Indiana Law Review

Volume 31 No. 1 1998

"LIES, DAMN LIES AND STATISTICS: HOW EMPIRICAL RESEARCH SHAPES HEALTH LAW AND POLICY"

A SYMPOSIUM IN CELEBRATION OF THE TENTH ANNIVERSARY OF THE CENTER FOR LAW AND HEALTH

Indiana Law Review

Volume 31 1997-98

Please enter my subscription to the INDIANA LAW REVIEW

NAME _____

ADDRESS _____

_____

_____

Enclosed is $_____ for ____ subscriptions.

Bill me for _____ subscriptions.

Mail to:

INDIANA LAW REVIEW
INDIANA UNIVERSITY
SCHOOL OF LAW—INDIANAPOLIS
735 West New York Street
Indianapolis, Indiana 46202

Subscription Rates (one year):

Regular, $25; Foreign, $28 Student, $17 (4 issues)
Single Issue, $8; Survey Issue, $17;
Centennial Issue, $12

INDIANA LAW REVIEW

(ISSN 0090-4198)

Published four times a year by Indiana University. Editorial and Business Offices are located at:

Indiana Law Review
735 W. New York Street
Indianapolis, IN 46202-5194
(317) 274-4440

Subscriptions. The current subscription rates are $25.00 per four-issue (domestic mailing) and $28.00 (foreign mailing). Unless the Business Office receives notice to the contrary, all subscriptions will be renewed automatically. *Address changes must be received at least one month prior to publication to ensure prompt delivery and must include old and new address and the proper zip code.*

Single and Back Issues. Current single regular issues may be obtained from the *Law Review* Business Office for $8.00. Current Survey issues may be purchased for $17.00. The Symposium Issue may be purchased for $12.00. Please enclose payment with order. Complete sets and single issues of prior volumes are also available from the *Law Review* Business Office. Unless a claim is made for nonreceipt of issues within one year of the mailing date, we cannot guarantee supplying the issues without charge.

Manuscripts. The *Indiana Law Review* invites submission of unsolicited manuscripts. All citations should conform to *The Bluebook: A Uniform System of Citation* (16th ed. 1996). Manuscripts should be double-spaced with standard one-inch margins. The *Indiana Law Review*, as a matter of policy, promotes the use of gender-neutral language, and alternate use of the female or male pronoun is intended to refer to both female and male.

The INDIANA LAW REVIEW (ISSN 0090-4198) is the property of Indiana University and is published quarterly by the Indiana University School of Law—Indianapolis, which assumes complete editorial responsibility thereof. Subscription rates: one year $25.00; foreign $28.00. Please notify us one month in advance of any change in address and include both old and new addresses with zip codes to ensure delivery of all issues. Send all correspondence to Editorial Assistant, *Indiana Law Review*, Indiana University School of Law—Indianapolis, 735 W. New York Street, Indianapolis, Indiana 46202-5194. Publication office: 735 W. New York Street, Indianapolis, Indiana 46202-5194. Periodicals postage paid at Indianapolis, Indiana 46201.

POSTMASTER: Send address changes to INDIANA LAW REVIEW, 735 W. New York Street, Indianapolis, Indiana 46202-5194.

The entire text of this Law Review is printed on recycled paper.

Indiana University School of Law—Indianapolis

1997-98 ADMINISTRATIVE OFFICERS AND FACULTY

Administrative Officers

MYLES BRAND, *President of the University.* Ph.D. in Philosophy, University of Rochester (New York).

GERALD L. BEPKO, *Vice-President for Long-Range Planning, Indiana University; Chancellor, Indiana University-Purdue University at Indianapolis and Professor of Law.* J.D, ITT/Chicago-Kent College of Law; LL.M., Yale University.

NORMAN LEFSTEIN, *Dean and Professor of Law.* LL.B., University of Illinois; LL.M., Georgetown University.

SUSANAH M. MEAD, *Associate Dean for Academic Affairs and Professor of Law.* B.A., Smith College; J.D., Indiana University—Indianapolis.

JONNA M. KANE, *Assistant Dean for External Affairs.* J.D., Indiana University—Indianapolis.

ELIZABETH L. DECOUX, *Assistant Dean for Student Affairs.* J.D., Mississippi College School of Law.

JONI D. THOMPSON, *Assistant Dean for Development.* M.A., Ball State University.

ANGELA M. ESPADA, *Assistant Dean for Admissions.* J.D., Indiana University—Bloomington.

KAREN MARKS, *Director of Career Services.* B.S., St. Joseph's College; M.S., Illinois State University.

Faculty

CYNTHIA M. ADAMS, *Lecturer in Law.* B.A., Kentucky Wesleyan College; J.D., Indiana University—Indianapolis.

THOMAS B. ALLINGTON, *Professor of Law.* B.S., J.D., University of Nebraska; LL.M., New York University.

EDWARD P. ARCHER, *Professor of Law.* B.M.E., Rensselaer Polytechnic Institute; J.D., LL.M., Georgetown University.

JAMES F. BAILEY, III, *Professor of Law and Director of Law Library.* A.B., J.D., M.A.L.S., University of Michigan.

GERALD L. BEPKO, *Vice President for Long-Range Planning, Chancellor, and Professor of Law.* B.S., Northern Illinois University; J.D., IIT/Chicago-Kent College of Law; LL.M., Yale University.

ROBERT BROOKINS, *Professor of Law. B.S., University of South Florida*; J.D., Ph.D., Cornell University.

DANIEL H. COLE, *Professor of Law.* A.B., Occidental College; A.M., University of Chicago; J.D., Lewis and Clark College; J.S.M., Stanford Law School.

JEFFREY O. COOPER, *Assistant Professor of Law.* A.B., Harvard University; J.D., University of Pennsylvania.

PAUL N. COX, *Professor of Law.* B.S., Utah State University; J.D., University of Utah; LL.M., University of Virginia.

KENNETH D. CREWS, *Associate Professor of Law and Library and Information Science.* B.A., Northwestern University; J.D., Washington University; M.L.S., University of California, Los Angeles; Ph.D., University of California, Los Angeles.

CLYDE HARRISON CROCKETT, *Professor of Law.* A.B., J.D., University of Texas; LL.M., University of London (The London School of Economics and Political Science).

GEORGE E. EDWARDS, *Associate Professor of Law and Director, Program in International Human Rights and Law.* B.A., North Carolina State University; J.D., Harvard Law School.

PAUL J. GALANTI, *Professor of Law.* A.B., Bowdoin College; J.D., University of Chicago.

HAROLD GREENBERG, *Professor of Law.* A.B., Temple University; J.D., University of Pennsylvania.

JEFFREY W. GROVE, *Professor of Law.* A.B., Juniata College; J.D., George Washington University.

FRANCES WATSON HARDY, *Clinical Professor of Law.* B.S., Ball State University; J.D., Indiana University—Indianapolis.

MICHAEL R. HEISE, *Assistant Professor of Law and Director, Program in Law and Education.* A.B., Stanford University; J.D., University of Chicago; Ph.D., Northwestern University.

W. WILLIAM HODES, *Professor of Law.* A.B., Harvard College; J.D., Rutgers, Newark.

LAWRENCE A. JEGEN, III, *Thomas F. Sheehan Professor of Tax Law and Policy.* A.B., Beloit College; J.D., M.B.A., The University of Michigan; LL.M., New York University.

HENRY C. KARLSON, *Professor of Law.* A.B., J.D., LL.M., University of Illinois.

WILLIAM ANDREW KERR, *Professor of Law.* A.B., West Virginia University; J.D., LL.M., Harvard University; B.S., Duke University.

ELEANOR D. KINNEY, *Professor of Law and Co-Director of the Center for Law and Health.* B.A., Duke University; M.A., University of Chicago; J.D., Duke University; M.P.H., University of North Carolina.

RONALD J. KROTOSZYNSKI, Jr., *Assistant Professor of Law.* B.A., M.A., Emory University; J.D., LL.M., Duke University.

NORMAN LEFSTEIN, *Dean and Professor of Law.* LL.B., University of Illinois; LL.M., Georgetown University.

LYNN A. MCDOWELL, *Clinical Professor of Law.* B.S Indiana University; J.D., Indiana University—Indianapolis.

DEBORAH MCGREGOR, *Lecturer in Law.* B.A., University of Evansville; J.D., Georgetown University.

PATRICIA MCMATH, *Visiting Lecturer in Law.* B.A., J.D., Indiana University—Indianapolis.

SUSANAH M. MEAD, *Associate Dean and Professor of Law.* B.A., Smith College; J.D., Indiana University—Indianapolis.

MARY H. MITCHELL, *Professor of Law.* A.B., Butler University; J.D., Cornell Law School.

JAMES P. NEHF, *Professor of Law.* B.A., Knox College; J.D., University of North Carolina.

THOMAS R. NEWBY, *Lecturer in Law.* A.B., J.D., Indiana University—Bloomington.

LESLYE A. OBIORA, *Assistant Professor of Law.* LL.B., University of Nigeria; LL.M., Yale Law School; J.S.D. Candidate, Stanford Law School.

DAVID ORENTLICHER, *Associate Professor of Law and Co-Director of the Center for Law and Health.* A.B., Brandeis University; J.D., M.D., Harvard College.

JOANNE ORR, *Clinical Professor of Law.* B.S., Indiana State University; J.D., California Western.

DAVID R. PAPKE, *R. Bruce Townsend Professor of Law and Professor of American Studies.* A.B., Harvard College; J.D., Yale University; M.A. in American Studies, Yale University; Ph.D. in American Studies, The University of Michigan.

H. KATHLEEN PATCHEL, *Associate Professor of Law.* A.B., Huntington College; J.D., University of North Carolina; LL.M., Yale University.

FRAN QUIGLEY, *Clinical Professor of Law.* B.A., Hanover College; M.A., Indiana University—Bloomington; J.D., Indiana University—Indianapolis.

FLORENCE W. ROISMAN, *Associate Professor of Law.* B.A., University of Connecticut; Ph.D., Johns Hopkins University; LL.B., Harvard Law School.

JOAN M. RUHTENBERG, *Director of Legal Writing.* B.A., Mississippi University for Women; J.D, Indiana University—Indianapolis.

ANDREW T. SOLOMON, *Lecturer in Law.* A.B., University of Michigan; J.D., Boston University.

E. GARY SPITKO, *Assistant Professor of Law.* A.B., Cornell University; J.D., Duke University Law School.

JUNE O. STARR, *Associate Professor of Law.* B.A., Smith College; M.A., Columbia University; Ph.D., University of California, Berkeley; M.S.L., Yale University; J.D., Stanford Law School.

KENNETH M. STROUD, *Professor of Law.* A.B., J.D., Indiana University—Bloomington.

JAMES W. TORKE, *Professor of Law.* B.S., J.D., University of Wisconsin.

JAMES PATRICK WHITE, *Professor of Law and Consultant on Legal Education to the American Bar Association.* A.B., University of Iowa; J.D., LL.M., George Washington University.

LAWRENCE P. WILKINS, *Professor of Law.* B.A., The Ohio State University; J.D., Capital University Law School; LL.M., University of Texas.

MARY T. WOLF, *Director of Clinical Programs.* B.A., Saint Xavier College; J.D., University of Iowa.

Emeriti

AGNES P. BARRETT, *Associate Professor Emerita.* B.S., J.D., Indiana University.

CLEON H. FOUST, *Professor Emeritus.* A.B., Wabash College; J.D., University of Arizona.

DAVID A. FUNK, *Professor of Law Emeritus.* A.B., College of Wooster; J.D., Case Western Reserve University; M.A., The Ohio State University; LL.M., Case Western Reserve University; LL.M., Columbia University.

HELEN P. GARFIELD, *Professor Emerita.* B.S.J., Northwestern University; J.D., University of Colorado.

WILLIAM F. HARVEY, *Carl M. Gray Professor of Law & Advocacy Emeritus.* A.B., University of Missouri; J.D., LL.M., Georgetown University.

MELVIN C. POLAND, *Cleon H. Foust Professor of Law Emeritus.* B.S., Kansas State University; LL.B., Washburn University; LL.M., University of Michigan.

RONALD W. POLSTON, *Professor Emeritus.* B.S., Eastern Illinois University; LL.B., University of Illinois.

Law Library Faculty

JAMES F. BAILEY, III, *Professor and Director of Law Library.* A.B., J.D., M.A.L.S., University of Michigan.

MINDE C. BROWNING, *Assistant Director for Readers' Services.* B.A., Western University; M.L.S., Indiana University; J.D., Indiana University School of Law—Indianapolis.

RICHARD HUMPHREY, *Reference Librarian.* A.A., Brewton-Parker Junior College; B.A., Georgia Southwestern College; M.L.S., University of Kentucky.

WENDELL E. JOHNTING, *Assistant Director for Technical Services.* A.B., Taylor University; M.L.S., Indiana University.

BRUCE KLEINSCHMIDT, *Reference Librarian.* B.A., Furman University; J.D., University of Louisville: M.L.S., University of North Texas.

CHRIS E. LONG, *Catalog Librarian.* B.A., Indiana University; M.A., Indiana University; M.L.S., Indiana University.

MAHNAZ K. MOSHFEGH, *Acquisition/Serials Librarian.* B.A., National University of Iran; M.S., Tehran University; M.A., Ball State University; M.L.S., Ph.D., Indiana University.

KIYOSHI OTSU, *Computer System Specialist.* A.A., Parkland College; A.B., M.S., C.A.S., University of Illinois.

Indiana Law Review

Volume 31 **1998** **Number 1**

TABLE OF CONTENTS

TRIBUTE

SYMPOSIUM

NOTES

TRIBUTE TO PROFESSOR WALTER W. KRIEGER

DEBRA A. FALENDER*

Walter W. Krieger and I arrived at the Indiana University School of Law—Indianapolis at nearly the same time—he in the fall of 1971 as a first-time professor of property and agency law, and I in the fall of 1972 as a first-year law student enrolled in his Property I course. At that time, neither of us had any notion that our paths would be intertwined for the next twenty-five years.

For the entire first week of Property class in the fall of 1972, we discussed property rights in foxes pursued by hunters across a wasteland. The second week, we studied finders and keepers, and discovered that losers need not weep; and by the third week, we were ready to learn about property interests in buried treasure, bank notes on a window ledge, and rings in the mud at the bottom of a pool in the south of England. Many things were perplexing to us during the first few weeks of law school, but one fact was absolutely clear to me and to everyone else in that Property class—Professor Krieger was a wonderful, caring person, and we liked him and respected him immensely.

Twenty-five years later, the same fact remains absolutely clear—Professor Krieger is a wonderful, caring person, and we (his students and colleagues) like him and respect him immensely. In recognition of his teaching ability, the students voted him the recipient of the Black Cane Outstanding Professor Award in 1987. Recently, in 1997, the faculty honored him with a dinner and accolades on the occasion of his retirement. The *Indiana Law Review*, in this retirement tribute, also celebrates Professor Krieger's quarter-century career as a member of the faculty of the Indiana University School of Law—Indianapolis, and I am pleased to play a small part in that celebration, because Walter Krieger is my very special friend.

Walter Krieger was born, raised, and educated in Kentucky. He received his B.A. in 1959 from Bellarmine College in Louisville, and his J.D. (cum laude) in 1962 from the University of Louisville, where he was student editor of the Journal of Family Law and a member of the Phi Kappa Phi Honor Society and The Brandeis Society.

After completing law school, Krieger served as Judge Advocate in the U.S. Navy in Virginia and Hawaii, and ultimately was certified as a Military Judge by the Department of the Navy in 1969. Also in 1969, he received an LL.M. (with highest honors) in Public International and Comparative Law from the George Washington University National Law Center. Subsequently, he served in Washington, D.C., as the Assistant Head of the Law of the Sea Division in the Office of the Judge Advocate General of the Navy, until he resigned his regular commission as Lieutenant Commander in the U.S. Navy Judge Advocate General's Corps to accept a position on our faculty in 1971.

After such interesting experiences as prosecuting and defending clients in special and general courts-martial in Oahu, and advising the U.S. State Department during bilateral fisheries negotiations with Poland and the Soviet

---

* Professor of Law, Indiana University School of Law—Indianapolis, 1976-1997. J.D., 1975, Indiana University School of Law—Indianapolis; A.B., 1970, Mount Holyoke College.

Union, one wonders what it was like to arrive in landlocked Indianapolis in 1971 to begin teaching Property Law to first-year law students. To Walter Krieger, it surely felt good. He had returned to his midwestern roots. One gets a good sense of this from the stories he shared with his students—not stories of naval malpractice investigations, military life, or Washington politics, but stories of Kentucky Ducks, Red-Eyed Hogs, and Three Little Orphans, stories that gave life to the concepts of precedent, stare decisis, and equity.

Between August 1971 and his retirement in December 1996, Walter Krieger taught agency law, military law and international law, but early on, the focus of his teaching, writing, and outside speaking was in the areas of property and probate law. In 1980, 1982, and then continuously from 1984 until his retirement, he wrote the annual survey of property law for the *Indiana Law Review*.[1] His survey pieces are a delight to read and reread. He often included a gentle, sometimes subtle commendation for a particularly well-written and tightly analyzed decision, and occasionally he could not resist a polite but not-so-subtle chiding when a court didn't get the analysis or the conclusion quite right. His consistent and thoughtful work in the annual surveys had an important positive effect on the course of Indiana property law.

Walter Krieger's special research focus was in landlord-tenant law, where he relentlessly prodded the courts and the legislature to revise and modernize Indiana law in accord with national trends. In a 1977 article, Krieger and his co-author, Michael Shurn, criticized the Indiana legislature for failing to enact modern landlord-tenant legislation, and then challenged the courts to assume the responsibility.[2] Later, in a 1986 article, Krieger lamented that neither the courts nor the legislature had risen to the task of injecting certainty and fairness into Indiana landlord-tenant law, and once again, he gently encouraged them to do so in the near future.[3]

Walter wrote bar review outlines and taught bar review courses in agency/partnership and trusts for many years, until his health began to get the better of him in the summer of 1994. He also regularly lectured to lawyers in continuing education seminars. He was very well-received as a speaker, because

---

1. Krieger's Annual Surveys of Indiana Property Law for 1980, 1982, and 1984 through 1994, appear at 14 IND. L. REV. 459 (1981); 16 IND. L. REV. 283 (1983); 18 IND. L. REV. 347 (1985); 19 IND. L. REV. 299 (1986); 20 IND. L. REV. 305 (1987); 21 IND. L. REV. 343 (1988); 22 IND. L. REV. 369 (1989); 23 IND. L. REV. 485 (1990); 24 IND. L. REV. 1065 (1991); 25 IND. L. REV. 1375 (1992); 26 IND. L. REV. 1113 (1993); 27 IND. L. REV. 1285 (1994); 28 IND. L. REV. 1041 (1995).

2. Walter W. Krieger & Michael A. Shurn, *Landlord-Tenant Law: Indiana at the Crossroads*, 10 IND. L. REV. 591 (1977).

3. Walter W. Krieger, *Housing Code Violations and Tenant Remedies*, 19 IND. L. REV. 299, 322-23 (1986): "The application of the illegal contract doctrine to housing code violations and the recognition of an implied warranty of habitability in residential leases have replaced the old no duty to repair rule and the doctrine of *caveat lessee*. Nevertheless, in Indiana, the limited number of judicial decisions and the lack of legislation have left the scope of these new tenant remedies and landlord obligations uncertain."

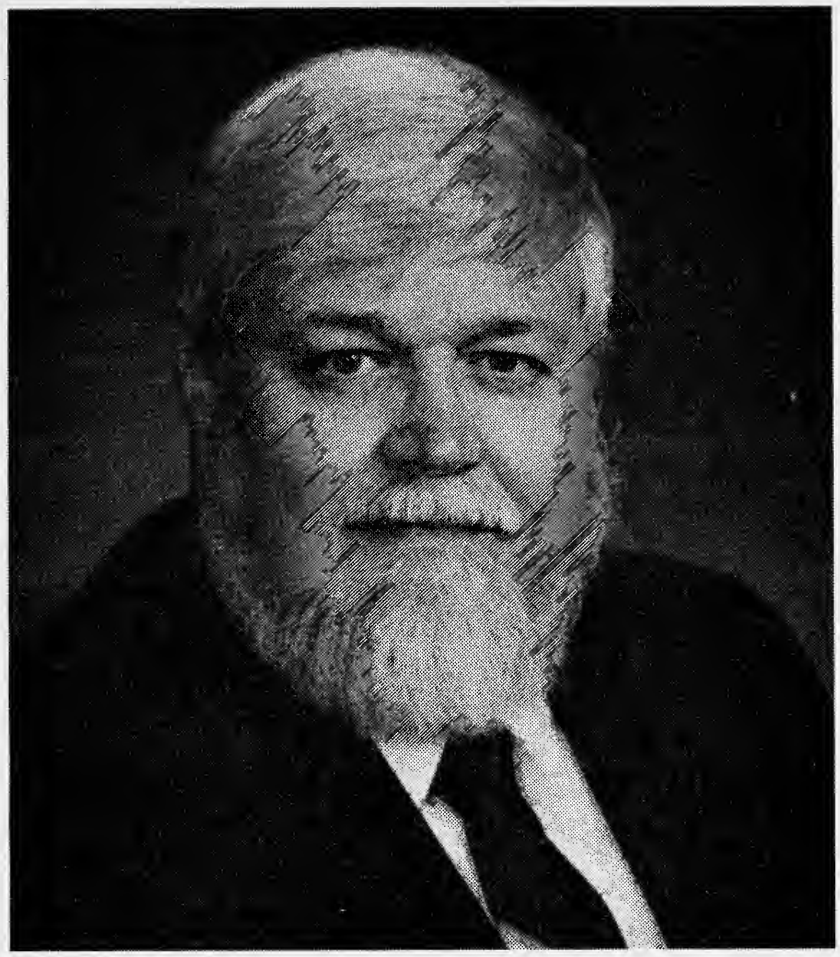
of his incredible knowledge of old and new cases and statutes, his delightful sense of humor, and his very pleasant demeanor, the very same traits that earned him the enduring respect of his law students over the years.

Walter always wore a smile, even in the midst of the most tedious committee meeting. He always said yes when asked to chair a committee, or judge a moot court competition, or present a model class to parents at Family Law Day. Every year, he received numerous thank-you letters from various law school and community organizations, for the generous devotion of his time and expertise.

Perhaps Walt said yes too often. Health problems plagued him in the last few semesters of his teaching. I know there were days when he had trouble just walking to his classroom. But he persevered, with a trademark twinkle in his eyes and an uncommon ability to share an amusing anecdote and make people smile.

Walter is a family man. He has a close and loving relationship with his lovely wife Carolyn, and I have never known a father more proud of a child's accomplishments and talents. In prominent locations in his office, Walter always displayed with pride the artwork of his son, James Edward.

Walter Krieger leaves the law school following a distinguished career. I and Walter's many close friends on the faculty wish Walter and his family all the best in the years to come.



WALTER W. KRIEGER, JR.

Education:

Bellarmine College, Louisville, B.A., 1959.
University of Louisville, J.D., 1962, *cum laude*.
The George Washington University, LL.M., 1969, *with highest honors*.

Academic Appointments:

1971-73 Assistant Professor of Law, Indiana University School of Law—Indianapolis.
1973-96 Associate Professor of Law, Indiana University School of Law—Indianapolis.

Honors and Awards:

Who's Who in American Colleges and Universities, 1959, 1962.
Student Editor, Journal of Family Law, 1961-62.
Phi Kappa Phi Honor Society, 1962.
Omicron Delta Kappa Outstanding Senior Award, 1962.
The Brandeis Society, University of Louisville, 1983.
Black Cane Award as Outstanding Professor, 1986-87 (Student Bar Association).

Subjects Taught:

Military Law; International Law; Agency; Property; Wills and Trusts; Fiduciary Administration; Future Interests.

Bar Admissions:
Indiana (1980).
Kentucky (1962).

Professional Experience:
Judge Advocate General Corps, United States Navy (active duty, 1962-71).
Staff Judge Advocate, U.S. Naval Hospital, Virginia, 1963-66 (Letter of Commendation, 1966).
Staff Judge Advocate, U.S. Naval Air Station, Hawaii, 1966-68 (additional duty Staff, Commander Fleet Air Hawaii).
Admitted to practice before the Court of Military Appeals, 1967.
U.S. Navy Post-Graduate Educational Program, The George Washington University, 1968-69.
Certified Military Judge of Courts-Martials of the Armed Forces, 1969.
Assistant Head of the Law of the Sea Division, International Law Division, Office of the Judge Advocate General of the Navy, Washington, D.C., 1969-71.
Department of Defense Advisor to U.S. State Department on Bilateral Fisheries Negotiations with (then) Soviet Union and Poland, 1970-71.

On faculty of the Indiana Continuing Legal Education Forum (ICLEF), in Seminars, Property and Trust areas, 1983-95.

Panelist in ICLEF seminars in the Property and Trust areas.

Bibliography:

Articles:
The United Nations Treaty Banning Nuclear Weapons and Other Weapons of Mass Destruction on the Ocean Floor, 3 J. MARITIME L. & COMM. 107 (1971).

Lieberman: How the Government Breaks the Law, 6 IND. L. REV. 629 (1973) (book review).

Walter W. Krieger, Jr. & Michael A. Shurn, *Landlord-Tenant Law: Indiana at the Crossroads*, 10 IND. L. REV. 591 (1977).

1980 Survey of Recent Developments in Indiana Law: Property, 14 IND. L. REV. 459 (1981).

1982 Survey of Recent Developments in Indiana Law: Property, 16 IND. L. REV. 283 (1983).

1984 Survey of Recent Developments in Indiana Law: Property, 18 IND. L. REV. 347 (1985).

Housing Code Violations and Tenant Remedies, 19 IND. L. REV. 299 (1986).

Developments in Property Law, 20 IND. L. REV. 305 (1987).

Property Survey 1987, 21 IND. L. REV. 343 (1988).

Survey of Indiana Property Law, 22 IND. L. REV. 369 (1989).

Survey of Recent Developments in Property Law, 23 IND. L. REV. 485 (1990).

Recent Developments in Property Law, 24 IND. L. REV. 1065 (1991).

Recent Developments in Property Law, 25 IND. L. REV. 1375 (1992).

1992 Developments in Indiana Property Law, 26 IND. L. REV. 1113 (1993).

1993 Developments in Indiana Property Law, 27 IND. L. REV. 1285 (1994).

1994 Developments in Property Law, 28 IND. L. REV. 1041 (1995).

Other Publications:

ICLEF manuals in Property and Trust areas.

Author of *Agency Outline*, *Partnership Outline*, and *Trust Outline* published in Indiana Bar Review Outlines by Professional Education Corporation. Produced teaching materials for course in Military Law taught in the mid 1970s. Participated from 1973-96 in Bar Review course.

Introduction:

"Lies, Damn Lies and Statistics: How Empirical Research Shapes Health Law and Policy"

A Symposium in Celebration of the Tenth Anniversary of the Center for Law and Health

ELEANOR D. KINNEY*
DAVID ORENTLICHER**

This symposium commemorates the tenth anniversary of the Center for Law and Health at Indiana University School of Law—Indianapolis. The Center was founded in 1986 as the Program for Law, Medicine and the Health Care Industry—the fulfillment of a longtime dream of many law school faculty and then Dean Gerald L. Bepko. The Board of Trustees established this program as a formal university center in June 1987. The program made great sense as nearly all the health professional schools of Indiana University, including the medical school, are located at the Indianapolis Campus. It is fitting that one of the first projects of the new Program was a symposium in the *Indiana Law Review* on "Financing and Regulating Health Care Services: Hard Choices and Ethical Dilemmas."

Chance and luck have much to do with the course of human events and the history of the Center for Law and Health is no different. In 1987, Professor Eleanor Kinney, the Center's founder and co-director, applied for and received a grant under The Robert Wood Johnson Foundation's Medical Malpractice Program to evaluate Indiana's Malpractice Act, an early and important effort at tort and insurance reform in a troubled area of tort law. With this project, the Center launched a commitment to empirical research.

Since 1987, the Center has engaged in many empirical studies addressing such topics as Medicare coverage policy, quality improvement for community-based, long-term care, barriers to health insurance for people with serious illness and designing strategies for expanding health coverage for the working poor. In addition, Professor Kinney has conducted many studies of grievance and appeal procedures as well as rule and policy making procedures for the Medicare and Medicaid programs with support from the Administrative Conference of the

---

* Professor of Law and Co-Director, Center for Law & Health, Indiana University School of Law—Indianapolis. B.A., 1969, Duke University; 1970, University of Chicago; J.D., 1973, Duke University School of Law; M.P.H., 1979, University of North Carolina at Chapel Hill.

** Associate Professor of Law and Co-Director, Center for Law & Health, Indiana University School of Law—Indianapolis. A.B., 1977, Brandeis University; M.D., 1981, J.D., 1986, Harvard University.

United States.

In 1995, Professor David Orentlicher joined the faculty at Indiana University School of Law—Indianapolis. Shortly thereafter, Professor Orentlicher became the co-director of the Center. With his national reputation as a excellent scholar in the field of bioethics, he brought a new dimension to the Center's work. His work has focused on the withdrawal of life-sustaining treatment, physician-assisted suicide, the effect of managed care on the patient-physician relationship, and rationing of health care.

The Center and its leadership are committed to the teaching mission of the law school. In addition to expanding the health law curriculum, the Center's leadership has sought to expand opportunities for students to learn more about health law and policy through internships, educational programs and employment opportunities to work on Center research projects. In the early 1990s, the Center worked with interested students to establish the student Health Law Society which organizes educational and social opportunities in the health care field. Annually, the Health Law Society and the Center sponsor a colloquium on important health law topics of current interest.

The paramount mission of the Center is research and scholarship. The Center's leadership shares a commitment that the Center's scholarship provide imaginative, practical and innovative ideas and approaches to the important problems facing the American health care system. Specifically, its scholarship should enhance access to high quality care for the most vulnerable of our society who are often excluded from the tremendous capabilities and services of the American health care system. It should also sort out the critical ethical dilemmas facing providers, payers and policy makers as they engage in the organization and delivery of health care services. Finally, the Center should serve as a resource to the State of Indiana as it faces new challenges in the financing, regulating and delivery of health care services.

Empirical scholarship has been one way in which the Center has distinguished itself historically. Thus, it is fitting that this symposium address the question of what impact empirical research has had and should have on health law and policy.

The subjects of the articles and commentaries all address areas of the Center's research both past and future. The Symposium opens with an article by Professor Sandra Johnson of Saint Louis University School of Law which reviews the empirical evidence on how end-of-life decisions are made and focuses, more importantly, on how that evidence is used by providers and policy makers to ensure better decisions in these trying circumstances.

The next three articles address empirical evidence and coverage policy. Professor William M. Sage of Columbia Law School analyzes judicial opinions on health insurance coverage and, specifically, the degree to which they can and should provide empirical evidence on coverage decision-making. Professor Maxwell Mehlman of Case Western Reserve University takes the analysis further in questioning what other credible empirical evidence on coverage decision-making besides case law is available. Karen A. Jordan, a professor of law at the

University of Louisville School of Law and a graduate of this law school, rounds out the discussion of the use of judicial decisions as empirical evidence by analyzing how empirical studies of judicial decisions affect the coverage policy making process.

Following are three articles on empirical evidence and antitrust law beginning with Professor James Blumstein's article on the interrelationship of empirical and normative issues in antitrust analysis. In the next article, a leading Indiana practitioner in health law, John C. Render of Hall Render Killian Heath & Lyman, comments on anomalies in the health market which should influence antitrust analysis in the health care field. The symposium concludes with an article from Professor Michael Jacobs of DePaul University School of Law who addresses how empirical evidence has recently been used in hospital merger analysis.

All articles make important contributions to legal scholarship and hopefully shed some light on the question of how empirical scholarship influences health law and policy. This inquiry on the role of empirical research on health law and policy addressed in this symposium will continue in an upcoming issue of the *Indiana Law Review* when Randall R. Bovjberg's article on empirical evidence and medical malpractice will appear in a separate symposium on Indiana's Medical Malpractice Act and the Supreme Court's consideration of four cases that challenge the constitutionality of that Act.

Indiana Law Review

Volume 31 1998 Number 1

ARTICLES

END-OF-LIFE DECISION MAKING: WHAT WE DON'T KNOW, WE MAKE UP; WHAT WE DO KNOW, WE IGNORE

SANDRA H. JOHNSON*

INTRODUCTION

Imagine that well-structured empirical studies consistently indicated that doctors do not tell patients what tests they are performing or why;[1] imagine that doctors can frame the information they provide patients and quite successfully generate the physician-desired consent or refusal of the treatment;[2] and imagine that only about half of patients recall being informed of serious risks of interventions, such as the risk of death.[3] Now, imagine one more thing: that outcome studies indicate that survival rates do not vary according to whether the physician informed the patient of significant risks, benefits and alternatives of treatment.[4]

---

* Professor of Law, St. Louis University School of Law. A.B., St. Louis University; J.D., New York University; LL.M., Yale Law School.

1. Lori B. Andrews, *Past as Prologue: Sobering Thoughts on Genetic Enthusiasm*, 27 SETON HALL L. REV. 893, 904 (1997) (describing practice of clinical genetic testing without disclosure to patients); Thomas Koenig & Michael Rustad, *His and Her Tort Reform: Gender Injustice in Disguise*, 70 WASH. L. REV. 1, 69-70 (1995) (detailing cases in which physician failed to inform female patients that certain surgical procedures would be done).

2. *See, e.g.*, Theresa M. Marteau, *Framing of Information: Its Influence Upon Decisions of Doctors and Patients*, 28 BRIT. J. SOC. PSYCHOL. 89, 93 (1989) (stating that options framed positively were more likely to be chosen than similar options framed negatively); David Orentlicher, *The Illusion of Patient Choice in End-of-Life Decisions*, 267 JAMA 2101 (1992) (stating that patient preferences regarding treatment options are often shaped by their physician's values); Cathy J. Jones, *Autonomy and Informed Consent in Medical Decisionmaking: Toward a New Self-Fulfilling Prophecy*, 47 WASH. & LEE L. REV. 379 (1990) (describing bias in information provided).

3. *See, e.g.*, Terence C. Wade, *Patients May Not Recall Disclosure of Risk of Death: Implications for Informed Consent*, 30 MED. SCI. L. 259, 260 (1990) (discussing a study of patients' failure to recall disclosure of the risk of death). *But see* Alan Meisel & Mark Kuczewski, *Legal and Ethical Myths About Informed Consent*, 156 ARCHIVES INTERNAL MED. 2521, 2524 (1996) (arguing that recall is not the equivalent of understanding and that understanding at the point of decision making is the critical issue).

4. Informed consent has not been proven to affect survival rates. However, it has been

If good empirical data revealed all of this about actual practice, what response would be most appropriate? Would we abandon the ethical and legal duty of informed consent as impractical and unworkable and require only that the doctor make treatment choices in the patient's best interest? Should we?

A serious challenge to the law and public policy supporting a requirement of informed consent has not yet materialized,[5] even though all but one of the data statements described above are authentic. Repeatedly, empirical studies have cast doubt on the practicality and performance of informed consent,[6] but the place of that value in law and ethics has not yet collapsed, and may never collapse, in response to the data.[7]

To what extent should data on actual practices determine public policy and legal and ethical standards in health care decision making? A first-level response to this question is whether the particular study has produced accurate and verifiable results and, assuming that the research was itself well-designed and that the data are solid, whether the results are being used within the boundaries and limitations of the study's design. All statistical evidence generates these questions. In fact, the conference at which the papers in this Symposium issue of the *Indiana Law Review* were originally presented was entitled *Lies, Damn Lies and Statistics?: How Empirical Research Shapes Health Law and Policy*.[8] The conference title quoted Mark Twain's identification of "three kinds of lies;"[9] but, in addition to providing a catchy title for a conference,[10] Twain's cutting

---

associated with other effects. *See, e.g.*, Wendy Levinson et al., *Physician-Patient Communication: The Relationship With Malpractice Claims Among Primary Care Physicians and Surgeons*, 277 JAMA 553 (1997) (discussing patient dissatisfaction and litigation resulting from breakdowns in physician-patient communications). In a substantial analysis of informed consent, Peter Schuck observed that studies on the costs of informed consent are lacking, but that there are reports of other positive outcomes including: achieving better results, decreased likelihood of malpractice claims, and enhanced perception of competence and control by the patient. Peter Schuck, *Rethinking Informed Consent*, 103 YALE L.J. 899, 943 (1994).

5. Schuck, *supra* note 4, at 902-03.

6. Meisel & Kuczewski, *supra* note 3, at 2521-22.

7. Schuck, *supra* note 4, at 902-03. In reference to this debate, Schuck identified three different versions of informed consent doctrine: the "letter and spirit of the doctrine" ("law in books"); the law as imagined or caricatured by some doctors ("law in the mind"); and the doctrine as actually practiced ("law in action"). *Id.* at 903. Schuck noted that commentators fall into two groups—the idealists and the realists—and that these groups talk past each other in discussions of informed consent. *Id.* at 903-04.

8. Symposium, *Lies, Damn Lies and Statistics?: How Empirical Research Shapes Health Law and Policy*, 31 IND. L. REV. 9 (1998).

9. MARK TWAIN, MARK TWAIN'S OWN AUTOBIOGRAPHY 185 (Univ. of Wis. Press 1990). In his autobiography, Twain quoted English politician and historian Benjamin Disraeli, who stated that "there [are] three kinds of lies: 'lies, damned lies, and statistics.'" Morningstar, Inc. v. Pilgrim Group, Inc., 29 Cal. Rptr. 2d 547, 553 (Cal. Ct. App. 1994).

10. Twain's original statement may have been closed with an ironic exclamation point, of course; surely not by a question mark. The addition of the exclamation is a slight offense, however,

critique of quantitative empirical research, at a minimum, identifies the skepticism and skill with which empirical research in bioethics should be analyzed.

This data-checking inquiry is not in itself an adequate response especially for any enterprise that sets standards of conduct, which includes both law and bioethics. Even well-designed statistically valid data requires interpretation.

When verifiable data on current practices clash with well-established normative standards, empirical research commonly appears to be appreciated only narrowly. The data may be construed, for example, to reveal only implementation or enforcement problems, while marginalizing the empirical challenge and leaving the core values and norms intact. Thus, it stimulates a call for more education, or greater "commitment," or stronger enforcement, rather than altering the normative standards.

In rare instances, empirical research may trigger radical calls for overthrowing the basic principles and paradigms that underlie established ethical and legal duties. One of the most notable occasions which prompted such a challenge occurred following investigation into actual practices in human experimentation, exposing evidence that medical researchers in the U.S. performed experiments on human beings in secret and without disclosure to their subjects.[11] In such a case, empirical research on actual behavior is seen as revealing an inadequacy in basic norms.[12]

As an applied ethics, bioethics must struggle with context. Empirical research tests whether ethical and legal standards "fit" the health care setting.[13] Bioethics, however, shares with law a fundamental tension in the essential norm-setting function between maintaining desired norms despite sometimes frequent violation and assuring that standards are realistic and rest on a substantial practice of voluntary compliance.

The controversy over the appropriate role of empirical research in law has persisted for over a century. Coincidentally, the Center's conference on

---

and the prominence accorded this commentary attests to the integrity of the groundbreaking empirical research conducted by the Indiana University School of Law—Indianapolis Center for Law and Health.

11. DAVID J. ROTHMAN, STRANGERS AT THE BEDSIDE: A HISTORY OF HOW LAW AND BIOETHICS TRANSFORMED MEDICAL DECISION MAKING 3 (1991).

12. In health law, these occasions arise in the unusual case in which the courts depart from the standard of the profession and establish a judicial standard of professional duty. The classic case is *Helling v. Carey*, 519 P.2d 981 (Wash. 1974). In informed consent doctrine, at least half of the states hold the physician to the standard of the profession rather than to the patient's expectations. BARRY R. FURROW ET AL., HEALTH LAW § 6-10(a) (1995). In cases involving the withdrawal of life-sustaining treatment, many of the earliest cases attempted to harmonize their stated legal rule allowing termination of treatment with the ethical standards of medical practice even though the practice of withdrawal was still controversial. *See, e.g.*, *In re* Quinlan, 355 A.2d 647, 664-69 (N.J. 1976).

13. Susan M. Wolf, *Shifting Paradigms in Bioethics and Health Law: The Rise of a New Pragmatism*, 20 AM. J.L. & MED. 395, 403-04, 409-10 (1994).

empirical research took place shortly after the centennial celebrations of the publication of Oliver Wendell Holmes' essay, *The Path of the Law*, in the *Harvard Law Review*.[14] In his essay, Holmes claimed: "For the rational study of the law, the black-letter man may be the man of the present, but the man of the future is the man of statistics . . . ."[15] Holmes was not merely predicting the influence of statistics,[16] which had Twain lamenting only a few decades later, he was speaking prescriptively. Holmes argued that the law *should* be designed so that its actual effects take precedence over arguments based solely on morality or history or philosophy.[17]

Holmes used the criminal law to illustrate his point that the impact of the law is the true test of the law,[18] but he might have used "law and bioethics" had it existed a century ago. Both fields, even more so than many other fields of law, are often viewed as embodying society's moral character,[19] or at least society's declaration of the correct ordering of civilized communities.[20] Both criminal law and bioethics bear the mark of notions of the good and the natural order of things.

Holmes said the following about crime and punishment:

> What have we better than a blind guess to show that the criminal law in its present form does more good than harm? . . . Does punishment deter?

---

14. Oliver Wendell Holmes, *The Path of the Law*, 10 HARV. L. REV. 457 (1897), *reprinted in* 110 HARV. L. REV. 991 (1997).

15. *Id.* at 1001.

16. Holmes' insight has proven largely true in the influence of several significant movements, including Legal Realism, Law and Society, Law and Economics, and Critical Legal Studies, which all emerge from an empirically-based analysis of the law. Each of these schools challenge what law is and what matters in law. Martha Minow, *The Path as Prologue*, 110 HARV. L. REV. 1023, 1023-25 (1997).

17. Robert W. Gordon, *The Path of the Lawyer*, 110 HARV. L. REV. 1013, 1013-14 (1997).

18. Holmes, *supra* note 14, at 1001-02.

19. *See* 136 CONG. REC. 12,251, at 12,260 (1990) (statement of Sen. Grassley) ("In a country that cherishes a separation between the state and any officially sanctioned religious practice, the criminal law is one of the few available institutions through which society can make a moral statement . . . ."); Sissela Bok, *At the Juncture of Theory and Practice: Remarks on Receiving the Henry Knowles Beecher Award*, 26 HASTINGS CTR. REP. 5 (May-June 1996); Thomas P. Griesa, *There Is No Case for Legalizing Drugs*, WALL ST. J., Aug 10, 1993, at A14 ("We should bear in mind that the foundational fact about the criminal law is that it is a moral judgment."); George F. Will, *The Sting of Shame*, WASH. POST, Feb. 1, 1996, at A21 ("[T]he criminal law's expressive function is to articulate society's moral condemnation.").

20. *See* 136 CONG. REC. 11,224, at 11,250 (1990) (statement of Sen. Hatch) ("Through the imposition of just punishment, civilized society expresses its outrage and sense of revulsion toward those who, by contravening its laws, have not only inflicted injury upon discrete individuals, but also weakened the bonds that hold communities together.") (quoting Stephen J. Markman & Paul G. Cassell, Comment, *Protecting the Innocent: A Response to the Bedau-Radclet Study*, 4 STAN. L. REV. 121, 157 (1988)); Edmund D. Pellegrino, *The Metamorphosis of Medical Ethics: A 30-Year Retrospective*, 269 JAMA 1158 (1993).

> Do we deal with criminals on proper principles? . . . If the typical criminal is a degenerate, . . . it is idle to talk of deterring him by the classical method of imprisonment. . . . If, on the other hand, crime, like normal human conduct, is mainly a matter of imitation, punishment fairly may be expected to help to keep it out of fashion.[21]

The calculation advocated by Holmes' century-old essay does not yet, and probably never will, determine the outcome of public controversies over criminal law. Arguments over whether the lack of deterrent effect is relevant to the use of the death penalty[22] and the effect of "three-strikes" sentencing[23] evidence the continuing struggle over the purpose of criminal law and punishment. Continuing disputes over the causes and prevention of crime also illustrate the intractability of answering Holmes' apparently empirical social inquiries.[24] Current debates in the law of bioethics have a similarly complex relationship with facts and values; rationality and irrationality; moral aspiration and consequentialist accounting; intuition and evidence.

The debate over the relative influence of aspiration and practice, though an old debate both in law and ethics, is being pressed now in bioethics because of the recent emergence of a body of substantial empirical research on the operation of bioethics in practice. The most extensive empirical examination of bioethics "at the bedside" to date is also the most recent. From 1989 to 1994, the Robert Wood Johnson Foundation funded an ambitious study of care for patients

---

21. Holmes, *supra* note 14, at 1002.

22. *See generally* MICHAEL L. RADELET & MARGARET VANDIVER, CAPITAL PUNISHMENT IN AMERICA (1988); Walter Berns et al., *The Death Penalty: A Philosophical and Theological Perspective*, 30 J. MARSHALL L. REV. 463 (1997); William J. Bowers et al., *A New Look at Public Opinion on Capital Punishment: What Citizens and Legislators Prefer*, 22 AM. J. CRIM. L. 77 (1994); Stephen B. Bright, *The Death Penalty as the Answer to Crime: Costly, Counterproductive and Corrupting*, 36 SANTA CLARA L. REV. 1069 (1996); Walter L. Gordon, III, *Death Penalty: National Disaster Visits California*, 33 SANTA CLARA L. REV. 283 (1993); Jordan Steiker, Essay, *The Long Road Up From Barbarism: Thurgood Marshall and the Death Penalty*, 71 TEX. L. REV. 1131 (1993).

23. *See generally* Meredith McClain, Note, *"Three Strikes and You're Out": The Solution to the Repeat Offender Problem?*, 20 SETON HALL LEGIS. J. 97 (1996); Mark W. Owens, Legislative Note, *California's Three Strikes Law: Desperate Times Require Desperate Measures—But Will It Work?*, 26 PAC. L.J. 881 (1995); Ilene M. Shinbein, Note, *"Three-Strikes and You're Out": A Good Political Slogan to Reduce Crime, But a Failure in Its Application*, 22 NEW ENG. J. ON CRIM. & CIV. CONFINEMENT 175 (1996); Victor S. Sze, Comment, *A Tale of Three Strikes: Slogan Triumphs Over Substance as Our Bumper-Sticker Mentality Comes Home to Roost*, 28 LOY. L.A. L. REV. 1047 (1995).

24. *See generally* Stephen Kinsella, *A Libertarian Theory of Punishment and Rights*, 30 LOY. L.A. L. REV. 607 (1997); Richard Lowell Nygaard, Essay, *On the Philosophy of Sentencing: Or, Why Punish?*, 5 WIDENER J. PUB. L. 237 (1996); Thomas J. Philipson & Richard A. Posner, *The Economic Epidemiology of Crime*, 39 J.L. & ECON. 405 (1996); Negley K. Teeters, *Fundamentals of Crime Prevention*, FED. PROBATION, Sept. 1995, at 63.

hospitalized with life-threatening conditions.[25] *SUPPORT*, which is discussed later in this Article, tested the actual operation of informed consent, patient autonomy, and compliance of physicians with patient or surrogate choice in medical treatment decision making, among other events in the treatment of these patients.[26]

SUPPORT contrasts with another type of influential study of bioethics, one with a longer history. In 1978, Congress established the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research.[27] The President's Commission studied bioethics from the position of logic and values and clinical experience, though largely intuitive or anecdotal experience and rarely empirical.[28] The President's Commission produced a document that provided the foundation for the development of the law and principles of bioethics regarding life-sustaining treatment decisions.[29] This document, *Deciding to Forego Life-Sustaining Treatment*,[30] has been cited as persuasive in no fewer than thirty-five appellate judicial opinions resolving end-of-life treatment issues.[31] In the nearly twenty years that passed between the President's Commission and *SUPPORT*, the basic principles of bioethics, especially as they were captured in law, remained grounded primarily in the "head work" of the Commission and its progeny.[32]

Although empirical research is a relative latecomer to bioethics, data on practices relating to medical ethics have already produced sometimes troubling questions concerning the gap between normative principles, including those adopted by the President's Commission, and reality. For example, are the basic values of American bioethics and the legal framework of health care decision making culture-bound and culturally exclusive of large numbers of U.S. patients?[33] Are surrogate decision makers, including family members, "reliable"

---

25. The SUPPORT Principal Investigators, *A Controlled Trial to Impede Care for Seriously Ill Hospitalized Patients: The Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatment (SUPPORT)*, 274 JAMA 1591 (1995) [hereinafter *SUPPORT*].

26. *Id.*

27. 42 U.S.C. § 300 (1978).

28. PRESIDENT'S COMMISSION FOR THE STUDY OF ETHICAL PROBLEMS IN MEDICINE AND BIOMEDICAL AND BEHAVIORAL RESEARCH, DECIDING TO FOREGO LIFE-SUSTAINING TREATMENT (Comm. Print 1983).

29. *Id.*

30. *Id.*

31. *See, e.g.*, Thor v. Superior Court, 855 P.2d 375, 381 (Cal. 1993); Care & Protection of Beth, 587 N.E.2d 1377, 1380 (Mass. 1992); *In re* Moorehouse, 593 A.2d 1256, 1261 (N.J. Super. Ct. App. Div. 1991); *In re* Fiori, 673 A.2d 905, 912 (Pa. 1996); *In re* Guardianship of L.W., 482 N.W.2d 60, 70-73 (Wis. 1992).

32. Similar efforts to guide the development of standards and practices in bioethics include much of the work of the Hastings Center, a private bioethics organization that has produced the influential Guidelines on the Termination of Life-Sustaining Treatment and the Care of the Dying (1987), among other documents, and The New York State Task Force on Life and Law.

33. *See, e.g.*, Leslie Blackhall et al., *Ethnicity and Attitudes Toward Patient Autonomy*, 274

in their understanding of the principal's preferences concerning medical treatment?[34] Even if a patient executes an advance directive, is the patient's doctor likely to comply with the document in directing medical treatment decisions?[35] Each of these questions has been studied empirically, and it is from such studies that an increasingly more robust body of empirical research in bioethics has formed.

There has not been a clear transition of authority in bioethics between the approach exemplified by the President's Commission of 1978 and the inquiry represented by the 1995 report of the *SUPPORT* results, however. Each form of bioethics "study," philosophical and empirical, continues to influence the area. Neither has achieved a commanding presence. Even increased and more powerful empirical research will not overtake the philosophical arm of bioethics, however, because the tension between the "is" and the "ought," between the real and the ideal, is inherent in any normative work that seeks to carry an authority in practice.

This Article examines how the debate over the appropriate source of legal and ethical norms in medicine has been played out publicly in recent judicial decisions regarding the legalization of physician-assisted suicide. It begins with an analysis of the Ninth Circuit's majority opinion in *Compassion in Dying v. Washington*.[36] It goes on to examine the results of *SUPPORT* and responses to those results. The Ninth Circuit opinion illustrates the use of empirical research and the framing of core questions in factual terms in efforts to change the law of bioethics. Responses to *SUPPORT* illustrate the persistence of basic principles in the face of substantially challenging empirical research.

I. *COMPASSION IN DYING V. WASHINGTON*

The Ninth and Second Circuit Courts of Appeals issued startling decisions in 1996. The courts' decisions in *Compassion in Dying v. Washington*[37] and in *Quill v. Vacco*[38] were startling simply because two federal appellate courts recognized a new constitutional right, an exceptional result in itself in light of the constitutional jurisprudence of this decade. The United States Supreme Court granted *certiorari* in both of those cases[39] and heard oral arguments in January

---

JAMA 820 (1995); Sheila T. Murphy et al., *Ethnicity and Advance Care Directives*, 24 J.L. MED. & ETHICS 108 (1996).

34. *See, e.g.*, P.M. Layde et al., *Surrogates' Predictions of Seriously Ill Patients' Resuscitation Preferences*, 4 ARCHIVES FAMILY MED. 518 (1995).

35. *See, e.g.*, Marion Danis et al., *A Prospective Study of Advance Directives for Life-Sustaining Care*, 324 NEW ENG. J. MED. 882 (1991).

36. 79 F.3d 790 (9th Cir. 1996), *rev'd sub nom.* Washington v. Glucksberg, 117 S. Ct. 2258 (1997).

37. *Id.*

38. 80 F.3d 716 (2d Cir. 1996), *rev'd*, 117 S. Ct. 2293 (1997).

39. *Compassion in Dying*, 79 F.3d at 790, *cert. granted sub nom.* Washington v. Glucksberg, 117 S. Ct. 37 (1996); *Quill*, 80 F.3d at 716, *cert. granted*, 117 S. Ct. 36 (1996).

1997. The course of the oral arguments foreshadowed the outcome of the cases and, as was thereafter generally expected, the Supreme Court reversed both circuit court decisions.[40] There was no dissent, but five justices joined in concurring opinions.[41]

The drama of a U.S. Supreme Court ruling on the legal status of assisted suicide concluded somewhat anti-climactically. The Court neither recognized a new or expanded right to die nor declared that the U.S. Constitution prohibited the states from recognizing or creating such a right within their own authority.[42] Justice Stevens' concurring opinion indicated he does not believe that the Court's decision closes the door on later consideration of a federal constitutional right to physician-assisted suicide under different circumstances.[43] Despite Stevens' opinion, the focus now moves to the states. State courts,[44] state legislatures,[45] and popular referenda/initiatives[46] are the arenas for the next round in the controversy over legalization of assisted suicide. As the arena shifts, the major points of argument reflected in these judicial opinions are likely to be repeated.

This Article focuses on the Ninth Circuit's majority opinion and only less so on the Supreme Court's majority opinion, although a note on the contrast between the two is appropriate. Although the Supreme Court authoritatively reversed the Ninth Circuit, the Supreme Court carried an easier burden than did the Ninth Circuit and than will the state courts and state legislatures. The Supreme Court did not take on the substantial burden of establishing a legal right in the face of great controversy. Nor did it have to justify institutionalizing the status quo under significant challenge. The Court's decision intentionally left the ultimate question of legalization open and would not close the door on the

---

40. *Glucksberg*, 117 S. Ct. at 2258; *Quill*, 117 S. Ct. at 2293.

41. *See Glucksberg*, 117 S. Ct. at 2275 (Souter, J., concurring in judgments); Washington v. Glucksberg, 117 S. Ct. 2302, 2303 (1997) (O'Connor, J., concurring in judgments); *Id.* at 2304 (Stevens, J., concurring in judgments); *Id.* at 2310 (Ginsburg, J., concurring in judgments); *Id.* (Breyer, J., concurring in judgments).

42. *Glucksberg*, 117 S. Ct. at 2258.

43. Justice Stevens specifically stated that "there are situations in which an interest in hastening death is legitimate . . . . I am also convinced that there are times when it is entitled to constitutional protection." *Glucksberg*, 117 S. Ct. at 2305. He remained open to the possibility that "an individual plaintiff seeking to hasten her death, or a doctor whose assistance was sought, could prevail in a more particularized challenge." *Id.* at 2309.

44. *See, e.g.*, Krischer v. McIver, 697 So. 2d 97 (Fla. 1997) (After the Supreme Court's decision, the Florida Supreme Court held that the Florida Constitution did not include a right to physician assistance in death.).

45. *See, e.g.*, S. 653, 89th Leg., Reg. Sess. (Mich. 1997); Henriette Campagne, *Lawmakers Take Up Assisted Suicide Issue*, MASS. LAW. WKLY., June 30, 1997, at 27.

46. Oregon is the only state that has approved the legalization of physician-assisted suicide, but the referendum vote is currently being challenged in the Oregon state courts. OR. REV. STAT. §§ 127.800-.897 (1996). The "Death with Dignity Act" will be sent back to the Oregon voters in November 1997. Richard Carelli, *Court Unanimously Rules Against Doctor-Assisted Suicides*, DET. FREE PRESS (June 27, 1997) <http://www.freep.com>.

matter.[47] Consistent with the jurisprudence of the Rehnquist Court, the Supreme Court simply endorsed the state's authority either to maintain or to change the law largely without constraint from the federal courts.[48] The Rehnquist opinion relied on historical state law prohibitions against aiding and abetting suicide[49] and accorded great deference to the state's assessment of potential harms under the rational basis test.[50] The Ninth Circuit, without the deference exercised by the Supreme Court, more substantively addressed arguments for and against legalization.[51] Therefore, the structure of its responses is most likely to reappear on the state level in the ensuing debate.

In addition to differing on the level of deference required, the Supreme Court's and the Ninth Circuit's majority opinions differed on the sources they relied upon as persuasive. To the extent that the Rehnquist opinion did engage the arguments for and against legalization substantively, it relied largely on the report of the New York State Task Force on Life and the Law.[52] In this, the Rehnquist opinion resembled state court decisions of the last two decades which relied heavily on the recommendations of the President's Commission to recognize a right to refuse or forego life-sustaining treatment.[53]

In contrast, in major parts of its majority opinion, the Ninth Circuit addressed the question of the legal status of assisted suicide with more of an emphasis on empirical evidence and its relevance for resolving public policy and legal issues. In other parts of its decision, the Court of Appeals made assumptions about professional behavior that make the question of legalization appear to rest on empirically verifiable facts. The Ninth Circuit's opinion is quite revealing in its treatment of arguments that rely on apparently empirically-based or empirically verifiable assertions, a point that is especially pertinent to the question presented by this special issue of the *Indiana Law Review*.

This critical analysis of the impact of "empiricalism" on the Ninth Circuit's decision is not intended to illustrate deficiencies unique to that opinion. An examination of the earlier Ninth Circuit opinion written for a panel of the Ninth Circuit by Judge John Noonan[54] and vacated by the Court of Appeals *en banc* in *Compassion in Dying*, for example, would likely yield a similar critique even though Judge Noonan's opinion reached a result opposite to that of the later majority. No matter the side, arguments for and against legalization of assisted suicide, to the extent they engage the substantive issues, present similar questions regarding the availability, reliability and relevance of empirical data on public and professional opinion, current practices, and anticipated risks and

---

47. Washington v. Glucksberg, 117 S. Ct. 2258 (1997).
48. *See id.* at 2275.
49. *Id.* at 2262-65.
50. *Id.* at 2271-75.
51. Compassion in Dying v. Washington, 79 F.3d 790, 816-39 (9th Cir. 1996).
52. *Glucksberg*, 117 S. Ct. at 2272-75.
53. *See supra* notes 27-32 and accompanying text.
54. Compassion in Dying v. Washington, 49 F.3d 586 (9th Cir. 1995), *vacated en banc*, 79 F.3d 790 (9th Cir. 1996), *rev'd sub nom.* Washington v. Glucksberg, 117 S. Ct. 2258 (1997).

benefits of the rejection or maintenance of the current legal status of assisted suicide.

The debate over physician-assisted suicide embraces many arguments. Three of several identifiable arguments in this debate are:

(1) The majority of Americans appear to indicate some support for physician-assisted death,[55] and therefore: law must change to reflect public opinion and actual practice;[56] *or* current law actually does accurately reflect society's "real" opinion concerning legalization,[57] if majority views are even relevant to the question.[58]

(2) The participation of doctors in the process of assistance in suicide, as compared to the participation of family members for example, alters the nature of the act: participation of doctors contributes to the safety of the process;[59] *or* the participation of doctors merely "medicalizes" or "white coats" the act so that it appears to be safer, more normal, and more acceptable.[60]

(3) Particular "vulnerable" populations may be affected differently by the legalization of assisted suicide: legalization will not harm these populations and may remedy the special suffering faced by vulnerable populations;[61] *or* legalization presents a special danger within a system that already victimizes certain identifiable populations.[62]

---

55. *See* James A. Tulsky et al., *A Middle Ground on Physician-Assisted Suicide*, 5 CAMBRIDGE Q. HEALTHCARE ETHICS 33, 34 (1995) (focusing also on the consistent refusal to indict, prosecute or convict physicians who aid in death); *see also* David Orentlicher, *Physician Participation in Assisted Suicide*, 262 JAMA 1844 (1989).

56. *See, e.g.*, Catherine L. Bjorck, *Physician-Assisted Suicide: Whose Life Is It Anyway*, 47 SMU L. REV. 371, 372 (1994) (arguing that opinion surveys, suicide statistics and public concern demand legalization of physician-assisted suicide).

57. Brief of the United States Catholic Conference et al. as Amici Curiae in Support of Petitioners at 8 n.1, Washington v. Glucksberg, 117 S. Ct. 2258 (1997) (No. 96-110); *see also* Stephanie Graboyes-Russo, *Too Costly to Live: The Moral Hazards of a Decision in* Washington v. Glucksberg *and* Vacco v. Quill, 51 U. MIAMI L. REV. 907, 915-16 (1997) (arguing that polls only reflect attitudes toward dying with dignity not toward physician-assisted suicide).

58. Edmund D. Pellegrino, *The Limitation of Empirical Research in Ethics*, 6 J. CLINICAL ETHICS 161, 162 (1995); *see also* Ruth Shalit, *When We Were Philosopher-Kings: The Rise of the Medical Ethicist*, NEW REPUBLIC, April 28, 1997, at 24 (criticizing bioethicists and the "place of ethical expertise in a liberal democracy").

59. *See, e.g.*, Lori D. Pritchard Clark, *Rx: Dosage of Legislative Reform to Accommodate Legalized Physician-Assisted Suicide*, 23 CAP. U. L. REV. 689, 705 (1994).

60. *See* John M. Dolan, *Is Physician-Assisted Suicide Possible?*, 35 DUQ. L. REV. 355, 392-93 (1996); *see also* George J. Annas, *Physician-Assisted Suicide—Michigan's Temporary Solution*, 20 OHIO N.U. L. REV. 561, 569 (1994) (arguing that physicians should not sacrifice their ethics to further the goals of the state).

61. Brief of Council for Secular Humanism and International Academy of Humanism as Amici Curiae in Support of Respondent at 9-10, Washington v. Glucksberg, 117 S. Ct. 2258 (1997) (No. 96-110).

62. *See, e.g.*, Alexander M. Capron, *Legalizing Physician-Aided Death*, 5 CAMBRIDGE Q.

Each of these three mirror-image points either relies on empirical data or suggests that empirical data will be relevant to the resolution of the dispute over the legalization of physician-assisted suicide. For example, the Ninth Circuit majority opinion in *Compassion in Dying* used public opinion polls as relevant evidence supporting judicial recognition of a constitutional right to physician-assisted suicide.[63] The court relied on the protective role of physicians as an essential characteristic of the legal right it recognized.[64] Further, the court rejected most claims that a legal right to assistance in suicide will operate in unacceptable population-specific patterns and argued that the availability of legalized physician assistance in suicide may address the particular needs of some persons within these groups.[65]

A. Opinion Polls

The Ninth Circuit described broad public support for the legalization of physician-assisted suicide. The court began its discussion of public opinion by observing that "[p]olls have repeatedly shown that a large majority of Americans—sometimes nearing 90%—fully endorse recent legal changes granting terminally ill patients, and sometimes their families, the prerogative to accelerate their death by refusing or terminating treatment."[66] The court further reported that "[o]ther polls indicate that a majority of Americans favor doctor-assisted suicide for the terminally ill,"[67] and described an April 1990 Roper poll that found "64% of Americans believed that the terminally ill should have the right to request and receive physician aid-in-dying."[68] The court also described "another national poll" that showed "nearly two out of three Americans favor doctor-assisted suicide and euthanasia for terminally ill patients who request it"[69] and stated that a 1994 Harris poll found "73% of Americans favor legalizing physician-assisted suicide."[70]

HEALTHCARE ETHICS 10, 19-20 (1996) (discussing the particular vulnerability women); Marshall B. Knapp, *Old Folks on the Slippery Slope: Elderly Patients and Physician-Assisted Suicide*, 35 DUQ. L. REV. 443, 444 (1996) (arguing that the elderly are a particulary vulnerable group susceptible to influence on physician-assisted suicide by family members and physicians because they may believe themselves a burden to their family or society).

63. *See* discussion *infra* Part I.A.

64. *See* discussion *infra* Part I.B.

65. *See* discussion *infra* Part I.C.

66. Compassion in Dying v. Washington, 79 F.3d 790, 810 (9th Cir. 1995) (citing Sanford H. Kadish, *Letting Patients Die: Legal and Moral Reflections*, 80 CAL. L. REV. 857, 860 & n.16 (1992)).

67. *Id.*

68. *Id.* (citing Robert L. Risely, *Voluntary Active Euthanasia: The Next Frontier, Impact on the Indigent*, 8 ISSUES IN L. & MED. 361, 365 (1992)).

69. *Id.* (quoting Kadish, *supra* note 66, at 861 n.22).

70. *Id.* (citing David Cannella, *Physician-Aided Suicide Fight Rages in Several States: Issue Expected to Go to Supreme Court*, ARIZ. REPUBLIC, May 13, 1995, at A22).

In the same paragraph with its recitation of the results of public opinion polls, the court also described the voting margins on the physician-assisted suicide referenda and state initiatives in Oregon, California, and Washington.[71] In the defendant state, Washington, the proposed legislation produced only 46% affirmative votes;[72] however, in Oregon, proposed legislation to legalize physician-assisted suicide had passed with 51% of voters approving the referendum.[73]

The court concluded its description of public opinion, including the results of the California, Washington and Oregon initiatives on legalization of assisted suicide, by concluding that "there is unquestionably growing popular support for permitting doctors to provide assistance to terminally ill patients who wish to hasten their deaths."[74] The opinion implied that the degree of public support for legalization is relevant to its resolution of the constitutional issue.[75] Growing approval could be used to indicate that its holding that there is a constitutional right to assistance in suicide reflects mainstream thought.

The Ninth Circuit's use of public opinion polls concerning physician-assisted suicide raises at least two issues that are specific to the suicide debate. First, although the Ninth Circuit considered public opinion to be relevant to the recognition of a liberty interest in physician assistance in death, it did not confront the gap between public opinion polls and the results of public votes on the issue to that point. Second, the court did not recognize that studies of both public and professional attitudes and practices relating to aid in dying are confounded by the transformation occurring in the language of end-of-life decisions.

The court described the results of the three public initiatives/referenda in the same paragraph as it discussed public opinion polls and concluded that the electoral results provide evidence of growing public support for legalization of assisted suicide. The court's reliance on the failed referendum in Washington state, the defendant in *Compassion in Dying*, as support for its holding that there is a constitutional right to assistance in suicide is curious. The court interpreted the clearly negative vote in Washington as positive support of its own position.[76]

The Ninth Circuit decision used evidence of public opinion polls to support its rejection of the results of a direct vote on the statute at issue in the case.[77] Public referenda may be irrelevant to any particular individual civil liberty,[78] of

---

71. *Id.*

72. *Id.*

73. *Id.*

74. *Id.* In November 1997, Oregon voters voted not to repeal their earlier affirmative vote on assisted suicide.

75. *Id.* at 811.

76. *Id.* at 810.

77. *Id.*

78. For a discussion of the role of popular votes on bioethical issues, see Judith F. Daar, *Direct Democracy and Bioethical Choices: Voting Life and Death at the Ballot Box*, 28 U. MICH. J.L. REFORM 799 (1995).

course, and the Ninth Circuit certainly did view access to physician-assisted death as a liberty interest that reached the most intimate decisions.[79] The Ninth Circuit, however, argued that public opinion polls were relevant to the recognition of this right. The court reviewed the jurisprudence on "fundamental rights" and noted the consistent effort to assure that the recognition of fundamental rights did not emerge only from the "imposition of the Justices' own choice of values."[80] The use of public opinion polls in constitutional adjudication is controversial. Generally, public opinion polls have not been persuasive when courts were dealing with issues of fundamental constitutional rights. Opinion polls have been labeled as being too uncertain because they are based on temporary opinions which are continually swayed by political activists.[81] The United States Supreme Court, considering a death penalty case, believed that opinion polls were inconclusive because the people had not voted on the issue through referendum or through their representatives in the legislature, which the court believed would be done if a position was truly popular.[82] Recently, a current Supreme Court Justice has spoken on the issue and submitted that the Federal Judiciary was given Article III protections (life tenure and salary protection) under the United States Constitution so that it could decide fundamental constitutional issues independent of such things as opinion polls.[83]

Despite its own discussion of public opinion, the Ninth Circuit finally specifically rejected both majority and minority control of this issue: "[N]either the state nor the majority of the people in a state can impose its will upon the individual" in this matter.[84] The court rejected arguments that the issue should be left to state elections, citing legal conflicts that may arise if some states legalized assisted death and others did not.[85]

Commentators have offered other reasons for rejecting the Washington and California votes as invalid indicators of public sentiment. For example, after the failed initiatives in California and Washington, commentators charged that the vote did not reflect the actual will of the majority because certain groups overwhelmed the proponents of legalization in terms of money spent on the campaign.[86] Others implied that certain "outside" groups heavily influenced the

---

79. *Compassion in Dying*, 79 F.3d at 812-13.

80. *Id.* at 803.

81. Swann v. Charlotte-Mecklenburg Bd. of Educ., 318 F. Supp. 786, 793 (W.D.N.C. 1970).

82. Stanford v. Kentucky, 492 U.S. 361, 377 (1989).

83. Hon. Justice Clarence Thomas, *Judging*, 45 U. KAN. L. REV. 1, 4 (1996) (Justice Thomas argued that if judges allow their "decisions to be guided by popular sentiment and group rights and demands, then the Constitution will be nothing but a malleable, transparent barrier to majoritarian desires.").

84. *Compassion in Dying*, 79 F.3d at 839.

85. *Id.* at 833 n.124.

86. Diane M. Gianelli, *Euthanasia Measures Fail, But Backers Vow to Renewed Push*, AM. MED. NEWS, Nov. 23, 1992, at 42; Susan Gilmore, *Will Foes' Efforts Doom California's Euthanasia Bill?*, SEATTLE TIMES, Oct. 24, 1992, at A14.

state-specific public debate.[87] Some advocates charged that opponents sensationally misrepresented the potential impact of the proposed legislation.[88] Others conceded that the legislation could have been rejected on the basis of specifics—the weakness of safeguards, for example—even though there may have been adequate support for more narrow legalization.[89] Still, some commentators have observed that pre-election polls may indicate that public support is strong enough to win as they did in Washington, but cannot indicate, until the actual vote, whether support is deep enough; that is, whether it will persist once arguments against the initially preferred position are offered or once an actual vote or other decision is required.[90]

The second issue in the court's use of survey data concerning acceptance of physician-assisted suicide relates to the emerging problem of language in life-sustaining treatment issues. The suicide debate, like some other politicized bioethics issues,[91] is plagued by jockeying over the names that will identify the advocate's position with virtue and caregiving and the opponent's position with evil and danger. In part, this is an unavoidable consequence of the politicization of the question: the battle takes place in snippets and slogans and has to appeal to emotion and intuition through association with familiar terms.

The language strategy may pay off politically. In 1993, in preparation for the referendum in Oregon, a survey was taken to test the impact of different terms.[92] In that survey, 44% of respondents indicated that they would vote for a law allowing "physician-aided suicide"; 51% for "physician's aid in dying"; 55% for

---

87. Jim Simon, *National Groups, Huge Budgets Invade Former Turf of the Little Guy*, SEATTLE TIMES, Oct. 4, 1991, at A1.

88. Ellis E. Conklin, *Support for Initiative 119 Slipped Away in Final Hours*, SEATTLE POST-INTELLIGENCER, Nov. 7, 1991, at A9; Tom Paulson, *Death With Dignity Backers Accuse Foes of Distortion*, SEATTLE POST-INTELLIGENCER, Oct. 24, 1991, at B1.

89. Howard Breuer, *Act of Mercy Brings Legal Problems Santa Paula Man Probed in Ailing Wife's Suicide*, L.A. DAILY NEWS, Oct. 17, 1993, at TO1; B.D. Colen, *Campaign '92 Initiatives California is Voting on the Right to Die Act Would Legalize Some Forms of Mercy Killing*, NEWSDAY, Oct. 29, 1992, at 23.

90. *See* Ezekiel J. Emanuel, *Empirical Studies on Euthanasia and Assisted Suicide*, 6 J. CLINICAL ETHICS 158 (1995); Gianelli, *supra* note 86, at 42; Conklin, *supra* note 88, at A9. The Ninth Circuit commented that the Washington and California proposals "contained far fewer practical safeguards" than did the "carefully-crafted" referendum in Oregon. *Compassion in Dying*, 79 F.3d at 810. Later in its opinion, however, the court rejected a central restriction in the Oregon referendum. The court implied in dictum that restricting assisted suicide to prescription of medication which the patient would administer to himself or herself may itself be unconstitutional. *Id.* at 831-32.

91. Another example is the debate over "partial-birth abortions." *See* Kathleen A. Cassidy Goodman, *The Mutation of Choice*, 28 ST. MARY'S L.J. 635 (1997); Kim Painter, *Handling the Aftermath of the Decision to Abort*, USA TODAY, Aug. 18, 1977, at 6D; Richard Saltus, *Late-term Abortion at Issue Called Rare in Mass.*, BOSTON GLOBE, May 17, 1997, at A1.

92. Mark O'Keefe, *Doctor-Assisted Suicide Law Boils Down to a War of Words*, THE OREGONIAN, Dec. 1, 1994, at A1.

"euthanasia"; and nearly 66% for a terminally ill patient's choice to "die with dignity."[93] The survey revealed an increase of 50% in affirmative responses as between "physician-aided suicide" and "death with dignity" just by use of the different term.[94] Equally revealing is the difference between "physician-aided suicide" and "physician's aid in dying":[95] the difference between 44% and 51% is the difference between winning and losing an election. The Ninth Circuit's use of polls asking about "aid in dying" as evidence of support for physician-assisted suicide demonstrates a lack of recognition of these language issues.[96]

Although the term "physician-assisted suicide" connotes professional assistance to an individual engaging in an act, committing suicide, that is not itself illegal, proponents of legalization no longer rely on the assisted "suicide" concept or language. The campaign director for the 1992 California initiative reacted to the campaign's victory in eliminating the word suicide from the ballot by stating that "the ballot language is worth a million bucks to us."[97]

The Oregon initiative itself provides that "Actions taken in accordance with this Act shall not . . . constitute suicide [or] assisted suicide."[98] The Ninth Circuit explicitly challenged the use of the "s" word to describe the intervention [99] even though the opinion laboriously examined the historical treatment of suicide to prove that suicide was not always viewed as immoral or illegal.[100]

The political power of language predictably has led advocates to search for a term for "assisted suicide" or "aid in dying" or "lethal intervention" that communicates a certain quality about the interaction of physician and patient. The language strategy will have an impact on the political process if politics holds true to form. But strategic language has produced a more substantial difficulty in this case. At a minimum, it has had an impact on the evaluation and meaning of empirical data describing current opinions and practices relating to end-of-life care. The poll discussed earlier,[101] which tested only differences in the terms, illustrates how powerfully the terminology can affect the results of surveys probing this issue.

The political use of language is not the only difficulty here. Underneath the political currency lies a fundamental deconstruction of the ethical and legal framework established for decisions regarding care at end of life. An essential point in the debate over assisted suicide is whether there is any difference at all between actions that are currently illegal and those that are legal and widely accepted. Advocates for legalization of assisted suicide argue that withholding

---

93. *Id.*

94. *Id.*

95. *Id.*

96. *Compassion in Dying*, 79 F.3d at 810.

97. Paul Jacobs, *Prop. 161—A Matter of Life or Death at the Polling Place*, L.A. TIMES, Oct. 10, 1992, at 20.

98. OR. REV. STAT. § 127.880 (1996) (Oregon Death with Dignity Act § 3.14).

99. *Compassion in Dying*, 79 F.3d at 824.

100. *Id.* at 806-10.

101. *See* O'Keefe, *supra* note 92.

or foregoing life sustaining treatment (particularly nutrition and hydration, but also ventilator support and other interventions) is no different than providing lethal medication.[102] Proponents of legalization also argue that administering medication to relieve pain, where the pain medication may foreseeably hasten the patient's death, is no different than providing a non-therapeutic lethal drug.[103] Withdrawal of medical treatment is legal within broadly circumscribed circumstances while providing or administering a lethal drug is not. Prescribing or administering pain medication therapeutically, even assuming that the medication may hasten death, which itself is a highly controversial point,[104] is not illegal.[105] Opponents of legalization argue that there is an ethical, historical, medical, and legal difference between aiding suicide and withdrawing treatment or treating pain.[106] The Ninth Circuit itself recognized that both withdrawal of treatment and provision of adequate pain relief are legal.[107] The Ninth Circuit adopted the argument that there is no legally significant difference between those actions and providing lethal medication, while the Supreme Court, in contrast, detailed the states' consistent adherence to legal distinctions between withdrawing treatment and medicating for pain on the one hand and providing lethal medication on the other.[108] The Rehnquist opinion and the concurring opinions further emphasized the legality of providing pain medication even where that medication may present the risk of an earlier death.[109]

Establishing a distinction in legal and ethical character between withdrawal of treatment and euthanasia or assisting suicide was a critical point in the courts' initial legal recognition of a patient's right to refuse medical treatment.[110]

---

102. *See, e.g.*, Brief of Julian W. Whitaker, M.D., as Amicus Curiae in Support of Respondents at 15, Washington v. Glucksberg, 117 S. Ct. 2258 (1997) (Nos. 96-110, 95-1858) [hereinafter Whitaker]; Brief of Gay Men's Health Crisis and Lambda Legal Defense and Education Fund as Amicus Curiae in Support of Respondents at 24, Washington v. Glucksberg, 117 S. Ct. 2258 (1997) (Nos. 96-110, 95-1858).

103. *See* Whitaker, *supra* note 102.

104. *See, e.g.*, John Colin Partridge & Stephen N. Wall, *Analgesia for Dying Infants Whose Life Support is Withdrawn or Withheld*, PEDIATRICS, Jan. 1, 1997, at 76 (reporting a study indicating that average length of time between withdrawal of ventilator support and death was identical for infants receiving no pain medication and those receiving morphine and longer for those infants receiving a greater amount of morphine).

105. Yale Kamisar, *The "Right to Die": On Drawing (and Erasing) Lines*, 35 DUQ. L. REV. 481, 500 (1996).

106. Brief Amicus Curiae in Support of Petitioners at 3, Vacco v. Quill, 117 S. Ct. 2293 (1997) (No. 95-1858); Brief of the United States Catholic Conference et al. as Amici Curiae in Support of Petitioners at 4-5, Vacco v. Quill, 117 S. Ct. 2293 (1997) (No. 95-1858).

107. Compassion in Dying v. Washington, 79 F.3d 790, 827 (9th Cir. 1996).

108. Washington v. Glucksberg, 117 S. Ct. 2258, 2270 (1997).

109. *Id.* at 2270; *id.* at 2276 (Souter, J., concurring); Washington v. Glucksberg, 117 S. Ct. 2302, 2307 (O'Connor, J., concurring).

110. *See, e.g.*, Alexander M. Capron, *Borrowed Lessons: The Role of Ethical Distinctions in Framing Law on Life-Sustaining Treatment*, 1984 ARIZ. ST. L.J. 647.

Persuading health care professionals that there was a qualitative ethical difference between allowing a patient to die and killing a patient was also important in the effort to assure that withdrawal of life-sustaining treatment and effective pain management would be carried out.[111]

Whether or not there is a "real" difference between withdrawing treatment or medicating for pain on the one hand and assisted suicide or euthanasia on the other, the suicide debate's deconstruction of the legal and ethical underpinnings of the last two decades (and along with that foundation, the language) has made it difficult to communicate. A prominent example of this difficulty emerged from a recent survey of critical care nurses.[112] This survey was described as revealing that nineteen percent of the nurses had reported performing or participating in "euthanasia," with the implication that the nurses had engaged in illegal or unethical behavior, that they had killed their patient.[113] This was not just another incident where the media could be criticized for misunderstanding or hyperbole. The study's principal investigator, Dr. David Asch, described his survey as supporting that finding.[114]

The publication of the survey elicited a strong response. Much of the negative reaction was heavily critical of the survey questions. Two questions in particular were controversial.

The survey asked whether the nurse respondent had ever performed actions "with the intent of causing or hastening that patient's death—other than the withdrawal of life-sustaining treatment" and whether the nurse had ever given an "overdose of opiates."[115] Both of these questions alluded to the administration of medication to a person whose death is imminent or who is in the final stages of terminal illness. In response to the publication of the Asch study, nurse specialists, professional nursing organizations, and nurse researchers generally claimed that the questions asked on the survey were too ambiguous to generate a meaningful response.[116]

The question referring to an "overdose of opiates" is particularly problematical. "Overdose" implies that the pain medication was administered in an amount that was not required for the treatment of pain but that was instead

---

111. Graboyes-Russo, *supra* note 57, at 916-17.

112. David A. Asch, *The Role of Critical Care Nurses in Euthanasia and Assisted Suicide*, 334 NEW ENG. J. MED. 1374 (1996).

113. Barbara Reynolds, *Nurses, Physicians Should Not Play God*, USA TODAY, June 7, 1996, at 15A; *Are Nurses Angels of Death?*, TIMES UNION, June 1, 1996, at A6.

114. David A. Asch, *Promises and Pitfalls Along the Road to Empirical Scholarship in Bioethics*, 2 CENTER FOR BIOETHICS NEWSLETTER 1 (1996).

115. David A. Asch, *The Role of Critical Care Nurses in Euthanasia and Assisted Suicide*, 335 NEW ENG. J. MED. 971, 973-74 (1996) (Letter to the Editor).

116. Colleen Scanlon, *Euthanasia and Nursing Practice—Right Question, Wrong Answer*, 334 NEW ENG. J. MED. 1401 (1996); *see also* Patricia A. Dunn, *The Role of Critical Care Nurses in Euthanasia and Assisted Suicide*, 335 NEW ENG. J. MED. 971 (1996) (Letter to the Editor); Nancy L. Szaflarski & John M. Clochesy, *The Role of Critical Care Nurses in Euthanasia and Assisted Suicide*, 335 NEW ENG. J. MED. 971 (1996) (Letter to the Editor).

non-therapeutic. Current research on effective pain relief has clearly established that "standard" doses of opiates are meaningless for patients who have received these pain medications over some amount of time.[117] Much larger doses of the drugs are both safe for such patients and are required for relief of pain.[118] Such "overdoses" can be well tolerated by those patients often without the sedation side effect or other side effects. For minimally adequate treatment of pain in patients who have used prescribed pain medication previously, doctors *should* be prescribing and nurses administering "overdoses" of opiates and can prescribe an "overdose" without shortening the patient's life.[119] If Asch intended to exclude these totally therapeutic, non-life-threatening "overdoses" from the reach of his question, he failed to do so. If he intended to include any case in which a "large" amount of opiates is provided, he has asked a question using a term that implies substandard care to describe acceptable medical or nursing care.

Asch responded to the criticism of his study:

> [T]he term "euthanasia" is loaded. Those who believe that the term can refer properly only to activities that are immoral may also feel that it cannot apply to all the activities reported by the nurses. If there is a continuum of moral appropriateness represented here, it is not clear where the moral divide lies, whether there is a single divide, or whether that divide is shifting over time. The range of activities described by the nurses who participated in this study may reveal the inadequacy of the term "euthanasia" and the many professional and legal policies built on it.[120]

Asch's response highlights the ambiguity of the term euthanasia but it also illustrates the difficulty of the context in which the debate over and studies of physician-assisted suicide are occurring. Whether or not the argument that there is no difference, other than the current legal distinction, between withdrawing treatment and medicating for pain on one hand and assisting in suicide on the other ultimately prevails, the "no distinction argument" has a direct impact on the meaningfulness of empirical studies on professional attitudes and practices. To the extent that such research may be used in judicial or legislative decision making, this problem has to be recognized specifically and with some sophistication.[121]

The "no distinction argument" could be having other broad and disturbing consequences. The desired outcome of the deconstruction of the current ethical and legal framework may be the expansion of medical options and the expansion of individual control over medical decisions, as well as the relief of suffering.

---

117. *See, e.g.*, Russell K. Portenoy, *Opioid Therapy for Chronic Nonmalignant Pain: Clinicians' Perspective*, 24 J.L. MED. & ETHICS 296 (1996).

118. *Id.*

119. *See, e.g.*, Szaflarski & Clochesy, *supra* note 116.

120. Asch, *supra* note 115, at 973.

121. For a critical analysis of survey data concerning physician attitudes toward euthanasia and assisted suicide, see Emanuel, *supra* note 90, at 158.

It is possible, however, that dismantling the legal and ethical framework for the treatment of pain could have an opposite effect. If medicating for pain relief becomes associated with physician-assisted suicide or euthanasia, adequate treatment for pain could take on all of the moral connotations of killing. Individual doctors and nurses who believe that it is immoral to "kill" a patient and who may already be anxious about providing adequate pain medication for their patients for fear of causing death, may become even more hesitant to provide medically required pain medication.[122] This reaction would exacerbate the current shameful neglect of pain and cause increased and avoidable human suffering. Furthermore, legislators who want either to make assisted suicide and euthanasia illegal, or to legalize assisted suicide and euthanasia but only under tightly controlled circumstances, might establish standards that greatly restrict access to effective pain relief. For example, in the 1997 legislative session, a bill was filed in the Michigan legislature that would legalize physician-assisted suicide.[123] A section of that bill, however, provided: "A nonphysician . . . who administers [or] causes to be administered . . . medications . . . to an individual for use in a manner substantially likely to cause or hasten that individual's death is guilty of a felony."[124] This provision would prohibit nurses from administering pain medication to patients who are suffering severe pain if the medication carries the risk of hastening death. Doctors prescribe medication, but nurses ordinarily administer medications in a hospital setting. Nurses ordinarily attend patients at the end of life. Physicians are often not readily available to administer pain medication as needed. This proposed legislation, then, legalizes the use of medication to assist a death when certain restrictive circumstances are met but could make the use of pain medication illegal where the narrow circumstances of the statute are not met.

B. How Physicians Do and Will Behave

Both *Vacco v. Quill* and *Compassion in Dying* limited their holdings to assistance provided by physicians. The three state initiatives placed before voters in California, Washington and Oregon also legalized assistance in suicide only where it was to be provided by physicians.[125]

Opponents, as well as some proponents, of legalization of assisted suicide have taken issue with the use of physicians to aid in death. Some opponents of legalization of assisted suicide argue that the involvement of physicians perverts the role of physician as healer and will irreparably damage the trust in the physician-patient relationship.[126] Other opponents of legalization argue that

---

122. Melissa L. Buchan & Susan W. Tolle, *Pain Relief for Dying Persons: Dealing with Physicians' Fears and Concerns*, 6 J. CLINICAL ETHICS 53, 57 (1995).

123. S. 653, 89th Leg., Reg. Sess. (Mich. 1997).

124. *Id.* § 5689.

125. Washington v. Glucksberg, 117 S. Ct. 2258, 2266 (1997); *see also* Compassion in Dying v. Washington, 79 F.3d 790, 810 (9th Cir. 1996).

126. Brief of the American Medical Association et al. as Amici Curiae in Support of

reliance on physicians as the centerpiece of the movement to legalize assistance in death is strategic to make the process appear to be an ordinary form of medical treatment.[127] Some proponents of legalization argue over whether doctors should have a monopoly over this work and whether other professionals, including nurses with prescribing authority or newly created categories of health care providers, would be better suited to the task.[128] Other proponents argue that the patient should be able to choose to have the assistance of family or friends instead of that of a professional stranger, and that legally requiring any professional intervention is an invasion of the privacy of the patient and gives too much control to physicians.[129]

The Ninth Circuit emphasized the protective and tempering role of physicians in the provision of assistance in suicide. It viewed the participation of doctors as a safeguard against abuse:

> We believe that most, if not all, doctors would not assist a terminally ill patient to hasten his death as long as there were any reasonable chance of alleviating the patient's suffering or enabling him to live under tolerable conditions. We also believe that physicians would not assist a patient to end his life if there were any significant doubt about the patient's true wishes. To do so would be contrary to the physicians' fundamental training, their conservative nature, and the ethics of their profession.[130]

In this statement, the court attributes two priorities to physicians in the care of their patients at the end of life. First, the court alludes to doctors' commitment to relieving suffering and pain. Second, the court emphasizes physicians' adherence to the wishes of their patients. The court nests these priorities in medical education and medical ethics.

No one can know how physicians will behave if assistance through lethal intervention were legalized. It would seem likely that neither Jack Kervorkian nor Timothy Quill—the two most prominent physician-advocates of physician-assisted suicide—would represent the majority of physicians providing such assistance. There is empirical evidence, however, that examines physicians'

Petitioners, Washington v. Glucksberg, 117 S. Ct. 2258 (1997) (No. 96-110); Brief of the Legal Center for Defense of Life, Inc. and the Pro-life Legal Defense Fund as Amici Curiae in Support of Petitioners, Washington v. Glucksberg, 117 S. Ct. 2258 (1997) (No. 96-110).

127. Brief of the American Medical Association et al. as Amici Curiae in Support of Petitioners, Vacco v. Quill, 117 S. Ct. 2293 (1997) (No. 95-1858); Brief of the American Medical Association et al. as Amici Curiae in Support of Petitioners, Washington v. Glucksberg, 117 S. Ct. 2258 (1997) (No. 96-110).

128. *See, e.g.*, Annette E. Clark, *Autonomy and Death*, 71 TULANE L. REV. 45 (1996).

129. David C. Thomasma, *An Analysis of Arguments For and Against Euthanasia and Assisted Suicide: Part One*, 5 CAMBRIDGE Q. HEALTHCARE ETHICS 62, 66-67 (1996) (describing claims that there should be no "middle person" involved and evidence of physician paternalism in refusals to comply with euthanasia requests).

130. *Compassion in Dying*, 79 F.3d at 827.

behavior in their care of patients generally, particularly in their treatment of patients in pain, and in their compliance with patients' choices concerning medical treatment.

The Ninth Circuit's confidence in the medical commitment to providing effective pain relief through available means is not supported in currently available data. The medical capacity to relieve pain is greater than it has ever been. Effective pain management is medically available for cancer pain, for chronic nonmalignant pain, and for pain related to diseases at the end of life. Effective pain management is available in many forms, including controlled substances such as opioids, that do not cause addiction or serious mental impairment.[131]

Yet, pain continues to be seriously neglected and undertreated. Treatable but untreated pain is a widespread problem that cuts across many patient populations. Studies have repeatedly documented undertreatment of pain in U.S. health care. For example, the *SUPPORT* study of 9105 patients dying in five teaching hospitals found pain management lacking. *SUPPORT* reported that the surviving family members of fifty percent of the dying patients reported that the patients suffered moderate to severe pain half of the time.[132] The very recent Institute of Medicine report on end-of-life care also identified undertreatment of pain as a major concern.[133] Seventy-five percent of cancer patients in one study reported suffering pain, with forty to fifty percent reporting moderate to severe pain and twenty-five to thirty percent reporting severe pain. This occurs even though ninety percent of cancer pain can be relieved through "relatively simple means."[134] Chronic nonmalignant pain has been described as an extremely prevalent problem.[135] Over two-thirds of nursing home residents experience serious chronic pain.[136] Moreover, the elderly, minorities, women, children, and those unable to speak for themselves due to disability bear the brunt of ineffective care and are undertreated at even higher rates than others.[137] Despite

---

131. *See* AGENCY FOR HEALTH CARE POLICY & RESEARCH, U.S. DEPT. OF HEALTH & HUMAN SERVICES, MANAGEMENT OF CANCER PAIN: MANAGEMENT OF CANCER PAIN GUIDELINE PANEL (1994) (Pub. No. 94-0592); AGENCY FOR HEALTH CARE POLICY & RESEARCH, U.S. DEPT. OF HEALTH & HUMAN SERVICES, ACUTE PAIN MANAGEMENT: OPERATIVE OR MEDICAL PROCEDURES AND TRAUMA (1992) (Pub. No. 92-0032).

132. *SUPPORT*, *supra* note 25, at 1591.

133. INSTITUTE OF MEDICINE, APPROACHING DEATH: IMPROVING CARE AT THE END OF LIFE (Marilyn J. Field & Christine K. Cassel eds., 1997).

134. AGENCY FOR HEALTH CARE POLICY & RESEARCH, *supra* note 131 (referring generally to both articles).

135. Russell K. Portenoy, *Opioid Therapy for Chronic Nonmalignant Pain: A Review of the Critical Issues*, 11 J. PAIN & SYMPTOM MGMT. 203 (1996).

136. Bruce A. Ferrell et al., *Pain in the Nursing Home*, 38 J. AM. GERIATRICS SOC. 409 (1990).

137. B. S. Shapiro & B. R. Ferrell, *Pain in Children and the Frail Elderly*, *in* APS BULLETIN (Oct./Nov. 1992); Knox H. Todd et al., *Ethnicity as a Risk Factor for Inadequate Emergency Department Analgesia*, 269 JAMA 1537 (1993).

the development of effective pain management interventions and the overall human and financial cost of pain, pain is neglected and undertreated.

The Ninth Circuit also expresses remarkable confidence in doctors' interest in and commitment to following the patient's wishes concerning medical treatment at the end of life. The Ninth Circuit's confidence greatly exceeds what the empirical data would support.[138]

In an important article published in the Journal of the American Medical Association, David Orentlicher, then writing for the Office of General Counsel of the AMA, and now co-director of Indiana University's Center for Law and Health, observes:

> Over the past two decades a societal consensus has developed around the principle that decisions about life-sustaining treatment should be guided by patient self-determination. According to the President's Commission, the Hastings Center, the American Medical Association, and the U.S. Supreme Court, treatment decisions should be based on the values, goals and preferences of the patient.
>
> While theory may emphasize the patient's values, empirical data suggest that other considerations may have a greater impact on decisions about life-sustaining treatment. *In particular, there is increasing evidence that physician values may be a more decisive factor than patient values in these decisions.*[139]

Orentlicher supports his statement with a review of the empirical literature. For example, the authors of a study of compliance with advance directives concluded that the physicians in the cases studied actually provided undesired treatment and withheld desired treatment based on the *physician's* judgment of what was appropriate or beneficial to the patient. Orentlicher concludes that "patient's preferences were respected as long as the physicians thought that the patients' choices resulted in the best decisions."[140]

Studies also consistently indicate that patients are greatly influenced by the manner in which the doctor presents treatment options. Orentlicher reviews studies indicating that patients are more likely to choose surgery when the probability of survival is presented rather than the probability of death,[141] and

---

138. *See* JAY KATZ, SILENT WORLD OF DOCTOR AND PATIENT (1984). Katz presents the classic critique of informed consent as a legal requirement and as a professional practice. In his analysis of the legal requirement of informed consent, Katz calls informed consent a "fairy tale" and delineates why informed consent is unnatural to physicians and how the law's enforcement of the requirement is ineffective and half-hearted at best.

139. David Orentlicher, *The Illusion of Patient Choice in End-of-Life Decisions*, 267 JAMA 2101, 2101 (1992) (emphasis added) (citations omitted).

140. *Id.* (discussing Marion Danis et al., *A Prospective Study of Advance Directives for Life-Sustaining Care*, 324 NEW ENG. J. MED. 882 (1991)).

141. *Id.* at 2102 (discussing T. M. Marteau, *Framing of Information: Its Influence Upon Decisions of Doctors and Patients*, 28 BR. J. SOC. PSYCH. 89 (1989)).

that patients are more likely to choose treatment when the treatment is presented positively rather than negatively.[142]

Orentlicher also reviews studies that indicate that physicians "are more inclined to talk with patients who are most like them"[143] and that physicians give both more time and more explanations of the course of treatment to patients who seem more intelligent and better educated.[144] Orentlicher concludes: "Ironically, physicians may become most aware of the preferences of patients who share their values. . . . Physicians may be less aware of the preferences of patients whose values diverge sharply from those of their physicians."[145]

More recent research remains consistent with the research relied upon by Orentlicher in his 1992 article.[146] In particular, the results of the *SUPPORT* study of in-hospital end-of-life care, discussed later in this article, confirms Orentlicher's conclusions.[147]

Empirical research on the medical neglect of pain and on non-compliance with patients' treatment choices at the end of life contradicts the Ninth Circuit's reliance on physicians as a safeguard in the administration of legalized physician-assisted suicide. The data actually support the proposition that physicians will not act as a guard against suicides undertaken because of neglected pain or without the complete understanding and consent of the patient.

Steven Miles calls for more study of the psychology of the physician-patient relationship in relation to physician-assisted suicide. He notes:

> There is ample evidence that physicians often have difficulty responding therapeutically to chronically ill or dying patients. They often underdiagnose and undertreat pain or depression, either of which may be part of why a patient may choose to die. Chronically ill or dying patients have told how their physicians emotionally or physically withdrew from them. . . .
>
> The intractability of these findings, despite years of calls for more sensitivity to the needs of dying persons and for better use of therapeutic information, implicates deep causes in the genesis of such mistreatment. . . . In this view, a physician tends to withdraw from dying patients, minimize pain, overlook clinical depression, and prematurely believe

---

142. *Id.* (discussing T.R. Malloy et al., *How Interventions Are Described Affects Patients' Decisions About Life-Sustaining Treatment*, *in* AMERICAN GERIATRIC SOCIETY/AMERICAN FEDERATION FOR AGING RESEARCH ANNUAL SCIENTIFIC MEETING, 1991 PROGRAM (Abstract A2)).

143. *Id.* at 2102.

144. *Id.*

145. *Id.*

146. *See, e.g.*, Lawrence J. Schneiderman et al., *Do Physician's Own Preferences for Life-Sustaining Treatment Influence Their Perceptions of Patients' Preferences? A Second Look*, 6 CAMBRIDGE Q. HEALTHCARE ETHICS 131 (1997); Steven H. Miles, *Physicians and Their Patients' Suicides*, 271 JAMA 1786 (1994).

147. *See* discussion *infra* Part II.A.

> patients have chosen to die as a result of the physician's own dread of death or disability or belief that a patient is as demoralized by his or her condition as the physician is. If this critique is correct, physicians are not bad, just human.[148]

Descriptions of current medical practices do not necessarily predict future practices. The patterns described by Orentlicher and *SUPPORT* and others might be remedied in the future. The Ninth Circuit relies on current medical training and ethics as a source of constraint for physicians in assisting death, and the court anticipates a future in which medical training and ethics will continue to operate in that fashion and will continue to emphasize constraints or prohibitions on intervening to cause death. Opponents to legalization argue that legalization itself will change medical ethics and medical education in the most significant way.[149] Perhaps these contrasting visions of the future could be empirically tested, although problems in structuring such a study seem insurmountable.

Even if such an experiment could be structured, it is unlikely that the data would resolve the question of whether the legalization of physician-assisted suicide has a positive or negative impact on physician behavior and medical ethics. Proponents of legalization argue that physicians who now cannot aid a patient in suicide without committing a crime, abandon their patients and their own ethical duty to relieve suffering.[150] In this view, the ethical physician is one who is willing to care for the patient even to the point of ending that patient's life, and a society that respects an individual's choice to end his life is a better society. Opponents argue that the physician's ethical duty to the patient requires that the physician be fully committed to the well-being of the patient, including aggressively treating pain even to the point of the patient's unconsciousness, but that the doctor's duty requires refusing the patient's request for death. From this perspective, a society that allows physicians to kill patients upon request is a society that has abandoned caring for the ill, the dying and dependent.[151] Empirical research could measure physicians' behavior pre- and post-legalization, but it cannot answer whether the behavior is more or less ethical than before.

The Ninth Circuit's confidence in physician practices in respecting their

---

148. Miles, *supra* note 146, at 1788-89.

149. Reply Brief for Petitioners, Washington v. Glucksberg, 117 S. Ct. 2258 (1997) (No. 96-110); Brief of Family Research Council as Amicus Curiae in Support of Petitioners, Washington v. Glucksberg, 117 S. Ct. 2258 (1997) (Nos. 96-110, 95-1858).

150. Brief of the American Medical Student Association and a Coalition of Distinguished Medical Professionals as Amici Curiae in Support of Respondents, Vacco v. Quill 117 S. Ct. 2293 (1997) (Nos. 95-1858, 96-110); Brief of Amicus Curiae Bioethicists Supporting Respondents, Vacco v. Quill 117 S. Ct. 2293 (1997) (Nos. 95-1858, 96-110).

151. Brief of the American Medical Association et al. as Amicus Curiae in Support of Petitioners, Vacco v. Quill, 117 S. Ct. 2293 (1997) (No. 95-1858); Brief of the International Anti-Euthanasia Task Force in Support of Petitioners, Vacco v. Quill, 117 S. Ct. 2293 (1997) (Nos. 95-1858, 96-110).

patients' wishes concerning medical treatment and in treating patients in pain is clearly misplaced. Part of the Ninth Circuit's argument in favor of a right to physician-assisted death, however, is that withdrawal of life-sustaining treatment under the current legal framework itself presents risks of abuse.[152] This observation could logically lead to restrictions on withdrawal of life-sustaining treatment just as much as it could lead to acceptance of the use of lethal drugs. The argument that there is no difference between the two can work both ways.

C. At-Risk Populations

Opponents to legalization of assisted suicide argue that legalization would place certain vulnerable populations at particular risk of involuntary, coerced, or simply desperate decisions for termination of life.[153] The groups that are often considered particularly vulnerable in the U.S. health care system include the poor, minorities, those disadvantaged by lack of access to adequate health care, the disabled and the elderly.[154] It is common practice for courts to identify particular groups as vulnerable and to construct legal rules designed to compensate for that vulnerability.

The Ninth Circuit reviewed arguments that these vulnerable populations will be at risk. However, the court of appeals rejected arguments that most of the vulnerable populations identified will be negatively affected by a change in the legal status of physician-assisted suicide. The court makes several different arguments on this point.

The court first addresses the argument that persons who do not have access to adequate health care will be threatened by the legalization of physician-assisted suicide:

> The argument that disadvantaged persons will receive more medical services than the remainder of the population in one, and only one area—assisted suicide—is ludicrous on its face. So, too, is the argument that the poor and the minorities will rush to volunteer for physician-assisted suicide because of their inability to secure adequate medical treatment.[155]

The Ninth Circuit is particularly harsh in its response to such arguments. It supports its conclusion with the experience of access to abortions, an instance where similar arguments about vulnerable populations have been made.[156] The court notes that public funding for abortions is very limited and that other factors

---

152. Compassion in Dying v. Washington, 79 F.3d 790, 817 (9th Cir. 1996).

153. Reply Brief for Petitioners Vacco and Pataki, Vacco v. Quill, 117 S. Ct. 2293 (1997) (No. 95-1858); Brief of the American Geriatrics Society as Amicus Curiae Urging Reversal of the Judgments Below, Vacco v. Quill, 117 S. Ct. 2293 (1997) (Nos. 95-1858, 96-110).

154. *Compassion in Dying*, 79 F.3d at 825.

155. *Id.*

156. *Id.*

converge to limit access even where there is an ability to pay.[157]

One could argue that the case of abortion services is not strictly analogous to this case. Abortion services are more costly and require more institutional collaboration, in terms of space and equipment, than would a prescription for lethal medication. Although abortion might not be a totally effective analogy, there may be other analogous situations.

Whether or not abortion is analogous to assisted suicide, a focus on the anticipated utilization rate reduces the argument over whether legalization specially endangers those who are currently denied health care and the poor and minorities to an empirically verifiable question. A study could be constructed to monitor the provision of assisted death services in a way that would test the court's assumption about excess utilization by such groups. Structuring the study would be difficult unless the use of the service could be both tightly controlled and accurately recorded, but it still might be feasible.[158]

Arguments based on inadequacies in the health care system include a further assertion that legalization of assisted suicide may undercut already weak public support for adequate health care, particularly for conditions that can trigger substantial costs with little or no hope for a return to health or substantially improved functioning. This concern reaches care for the dying and long-term care for the seriously debilitated elderly, among other situations.[159] This view argues that if the sometimes drawn-out "natural dying" process is viewed as a matter of personal choice, those who choose natural dying will have to bear the cost of that choice. The proposed regulation of assisted suicide in the Oregon initiative of 1994 attempts to respond to this problem and prohibits

> [t]he sale, procurement, or issuance of any life, health, or accident insurance or annuity policy or the rate charged for any policy shall not be conditioned upon or affected by the making or rescinding of a request by a person, for medication to end his or her life in a humane and dignified manner.[160]

Such regulation, however, cannot control the general level of insurance coverage or government payment for health care services for the dying, disabled or very dependent elderly.

Dissipation of public and private support for health care and supportive care for the dying and the debilitated elderly would be quite hard to measure. Such a change may operate glacially, and not be detectable for many years. This concern over withdrawal of financial support—whether public or private—is speculative, but so are assertions that legalization of assisted suicide will not

---

157. *Id.*

158. *See infra* text accompanying notes 164-65 (discussing various approaches to fairness).

159. *See, e.g.*, Susan Kastner, *The Battle Between Generations Many of Today's Sandwich Generation, the Newly Naked Middle Class, Have Been Forced Back to Frequently Bitter Interdependence. Young Against Old, Old Against Young, Families Cloistered Together, and at Odds*, TORONTO STAR, April 22, 1996, at A15.

160. OR. REV. STAT. § 127.875 (1996).

have such an impact.

The Ninth Circuit also addresses concerns about discrimination against the disabled. At one point, the court opined that such discrimination is unlikely because "[o]rganizations representing the physically impaired are sufficiently active politically and sufficiently vigilant that they would soon put a halt to any effort to employ assisted suicide in a manner that affected their clients *unfairly*."[161] The court did not describe what it meant by unfair treatment.

The situation of the disabled quite clearly raises the issue of what constitutes discrimination against any population. Is it discriminatory to legalize assisted suicide by disabled persons; or is it discriminatory to "protect" such individuals by excluding them from choosing certain options any other individual may choose? The Ninth Circuit in *Compassion in Dying* responded in this fashion:

> [s]eriously impaired individuals will, along with non-impaired individuals, be the beneficiaries of the liberty interest asserted here—and . . . if they are not afforded the option to control their own fate, they like many others will be compelled, against their will, to endure unusual and protracted suffering.[162]

The court's perception of assisted suicide as beneficial is a value judgment and is unlikely to be resolved empirically.

The Ninth Circuit offered fairness as a basis on which to resolve the conflict between risks to vulnerable populations generally and limitations on individual liberty: "The resolution that would be best for all, of course, would be to ensure that the practice of assisted suicide is conducted fairly and well, and that adequate safeguards sufficient to avoid the feared abuses are adopted and enforced."[163] In this argument, legalization would allow regulation of assisted suicide and activate safeguards to prevent abuses.

The aspiration that the system be administered "fairly and well" appears to be a standard that can be performance-monitored. One common and well accepted method for monitoring the fairness of any allocation system is to examine patterns in utilization. In recent years, for example, many studies on health care delivery have produced statistics on the provision of certain medical interventions across populations.[164]

Empirical studies on the incidence of assisted suicide across identified populations must operate on implied values of fairness: for example, that a

---

161. *Compassion in Dying*, 79 F.3d at 825 (emphasis added).

162. *Id.*

163. *Id.* at 825.

164. *See, e.g.*, M.B. Hamel et al., *Seriously Ill Hospitalized Adults: Do We Spend Less on Older Patients?*, 44 J. AM. GERIATRICS SOC. 1043 (1996); Scott Burris, *Dental Discrimination Against the HIV-Infected: Empirical Data, Law and Public Policy*, 13 YALE J. ON REG. 1 (1996); Carol Jonann Bess, *Gender Bias in Health Care: A Life or Death Issue for Women with Coronary Heart Disease*, 6 HASTINGS WOMEN'S L.J. 41 (1995); David R. Williams & Chiguita Collins, *U.S. Socioeconomic and Racial Differences in Health: Patterns and Explanations*, 21 ANN. REV. SOC. 349 (1995).

distribution that is determined solely by race or gender would be "fair" or "unfair." Empirical data standing alone cannot resolve the question of whether the system for assistance in death is operating "fairly." The resolution of whether assisted death is administered fairly requires that the content of the notion of fairness be made more specific. If fairness is taken in its most common meaning of "treating like cases alike," then the criteria for relevant similarities and differences need to be identified to provide the Ninth Circuit's notion of fairness with some content.

It is not clear from the Ninth Circuit's opinion, for example, how data that indicate a disproportionate representation of "seriously impaired individuals" among those choosing assisted suicide should be assessed. Would the data be interpreted as showing that assisted death was being administered fairly because it was reaching the people who could significantly benefit from the service, or would the data provide prima facie evidence of unfairness because of concerns that the disabled were not receiving adequate health care? The court's observation that "seriously impaired individuals" will be among the beneficiaries of legalization indicates that the court would not view disproportionate representation of disabled persons among those receiving medical services to end their lives as a problem.

The influence of economic factors in the administration of physician-assisted suicide might also be studied. Hypothetically, data could indicate either that patients receiving physician aid in death were primarily from higher income groups or that uninsured patients were disproportionately represented. If assisted suicide patients were often uninsured, would that indicate unfairness in the system? If only higher income patients were using this service, would it indicate that there were financial or non-financial barriers, such as required consultations and psychiatric assessments, that unfairly excluded lower-income uninsured patients? The Ninth Circuit opinion directly addressed the issue of financial considerations. In response to arguments that the terminally ill might choose assisted suicide because of financial pressures from "astronomical medical bills" that "consume the life savings they planned to leave for their families" or that "worse yet, burden their families with debts they may never be able to satisfy," the court stated that "[w]e are reluctant to say that, in a society in which the costs of protracted health care can be so exorbitant, it is improper for competent, terminally ill adults to take the economic welfare of their families and loved ones into consideration."[165] It is fair to say that the Ninth Circuit would not consider a disproportionate representation of persons of lower economic status an indication of unfairness in the administration of physician-assisted death. In fact, the court, in rejecting fears that the poor and minorities would be at risk of abuse, stated that there "is far more reason to raise the opposite concern: . . . that the poor and the minorities . . . will not be afforded a fair opportunity to obtain . . . the assistance that would allow them to end their lives with a measure of dignity."[166]

---

165. *Compassion in Dying*, 79 F.3d at 826.

166. *Id.* at 825.

The court did not address race-based patterns separately from income-based patterns, although the two can operate separately on access to health care.[167] Nor did the court separately address gender, although gender differences in health care have been discussed in the context of life-sustaining treatment decisions and physician-assisted suicide.[168]

II. THE STUDY TO UNDERSTAND PROGNOSES AND PREFERENCES FOR OUTCOMES AND RISKS OF TREATMENTS (*SUPPORT*)

If money were no object, could we design an empirical study that would determine once and for all whether we should maintain our allegiance to the ideal of patient choice in medical treatment decision making at the end of life? And if we could design such a study, would it resolve the issue of whether patient choice will drive physician-assisted death or whether physicians will drive physician-assisted death?

The Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments,[169] commonly known as *SUPPORT*, was a very large and very ambitious empirical study of the principles of bioethics in practice and certainly the most massive to date. *SUPPORT*, funded by the Robert Wood Johnson Foundation for $28 million,[170] studied the care of 9,105 patients admitted to five major medical centers in five states from June 1989 through January 1994. Eight co-principal investigators, with the advice of a ten-member national advisory committee, directed the efforts of over 150 professional staff engaged in the project.

Each of the over 9,000 patients was admitted to a participating hospital in the advanced stages of one of nine life-threatening or terminal illnesses. *SUPPORT* studied the care these patients received while in the hospital with a particular focus on resource utilization and on communication of and compliance with patient preferences. In their principal report on the project,[171] the investigators clearly view the importance of the study in relation to the basic principals of bioethics, especially as they have come to be understood by clinicians and institutionalized by the law in end-of-life care. The investigators note that "professional organizations, the judiciary, consumer organizations, and a

---

167. *See, e.g.*, Council on Ethical and Judicial Affairs, American Medical Association, *Black-White Disparities in Health Care*, 263 JAMA 2344 (1990); Vernellia R. Randall, *Racist Health Care: Reforming an Unjust Health Care System to Meet the Needs of African-Americans*, 3 HEALTH MATRIX 127 (1993).

168. Steven H. Miles & Allison August, *Courts, Gender and "The Right to Die,"* 18 L. MED. & HEALTH CARE 85 (1990); Nancy J. Osgood & Susan A. Eisenhandler, *Gender and Assisted and Acquiescent Suicide: A Suicidologist's Perspective*, 9 ISSUES L. & MED. 361 (1994).

169. *SUPPORT*, *supra* note 25, at 1591.

170. Steven A. Schroeder, Robert Wood Johnson Foundation, *President's Message, On Dying in America*, <http://www.rwjf.org/library/annual/pressr1.htm>.

171. The study produced many articles, but the principal report was published in the Journal of the American Medical Association. *See SUPPORT*, *supra* note 25.

president's commission have all advocated more emphasis on realistically forecasting outcomes of life-sustaining treatment and on improved communication between physician and patient."[172] The investigators view *SUPPORT* as an empirical test of the operation of informed consent and compliance with patients' preferences. In addition, the study associated patient choice with the issue of resource consumption in hospital care for patients near the end of life.[173]

The first years of *SUPPORT*, now known as Phase I, were conducted from June 1989 to June 1991. According to the investigators, Phase I revealed "shortfalls in patient-physician communication."[174] For example, of the 4,301 patients enrolled in the protocol in those years, 31% preferred that cardio-pulmonary resuscitation be withheld in the case of cardiac arrest, but fewer than half of the physicians treating these particular patients knew of their patients' preference to forego resuscitation.[175] For only 47% of all of the patients enrolled in Phase I could their physicians accurately report the patient's choice whether to accept or refuse resuscitation. Somewhat fewer than half of those patients who preferred not to undergo CPR had a written order (a "do-not-resuscitate order" or "DNR") which would allow CPR to be foregone should they arrest.[176] Because most hospitals require that CPR be performed unless there is a written order otherwise, the patients without such an order may have undergone unwanted resuscitative interventions if they had suffered an arrest.[177]

Pain was undertreated. Twenty-two percent of the patients interviewed reported suffering moderate to severe pain at least half the time in the hospital. Surrogates or family members interviewed after a patient had died reported that half of all conscious patients who had died in the hospital experienced moderate to severe pain at least half of the time over the last three days of hospitalization.[178]

The investigators reported that there was a substantial variation in these results among physician specialty groups and among the five medical centers. For example, the percentage of patients reporting moderate to severe pain ranged from a low of 12% in one institution to a high of 32% in another. Agreement on reports of DNR preferences varied by specialty from 8% for cardiologists and their patients to 24% for oncologists and their patients.[179]

The study was monitored throughout its course. Early findings revealed these substantial problems in physician-patient communication, treatment decisions and outcomes. When doctors at the participating hospitals indicated

---

172. *Id.* at 1591.

173. *Id.* The study used total hospital resource consumption per case and the number of days spent in intensive care when the patient was on ventilator support or in a coma as indicators. *Id.*

174. *Id.* at 1592.

175. *Id.* at 1594.

176. *Id.*

177. *Id.*

178. *Id.*

179. *Id.*

an interest in trying to change the situation, the study was altered in an attempt to change these observed patterns at the participating hospitals. Meetings were held with the doctors and other participants to develop a plan for improving physician-patient and family communication and for bringing treatment decisions, such as the decision concerning resuscitation, in alignment with patient preferences and probable outcomes. The result of these meetings became Phase II of *SUPPORT*.[180]

According to investigators, physicians participating in Phase I asserted that communication would improve if they could have access to more reliable and timely information and if project staff "would make it more efficient to have conversations."[181] In response to these ideas, the project designed an intervention to respond to the doctors' expressed needs. In Phase II of the project, patients were divided into a control group and an intervention group.[182] Patients and physicians in the intervention group received services from a *SUPPORT* nurse specifically focusing on providing information to the patient and family concerning treatment and prognosis and eliciting discussions with them concerning the patient's or surrogate's choices for treatment.[183] The patient's physician had to approve before the *SUPPORT* nurse could work with any particular patient.[184]

SUPPORT nurses worked with nearly all of the 2,652 patients enrolled in the intervention group of Phase II through a randomized process following hospital admission. Only 133 patients enrolled in the intervention group did not receive the *SUPPORT* intervention, and of these, 75 had died or were discharged on the day they were placed in the study.[185] The *SUPPORT* nurse intervened as seemed advisable in each case. The *SUPPORT* nurse talked directly with the patient or family about the patient's prognosis in 84% of the cases; about pain in 77%; about likely outcomes of resuscitation in 63%; and about written advance directives in 73% of the cases.[186]

A *SUPPORT* nurse discussed the patient's choices and understanding of their medical condition with the patient's doctor in "virtually all cases." At least one written report of the patient's or surrogate's preferences was provided in 78% of the cases.[187] *SUPPORT* nurses also engaged in "time-consuming discussions, arranged meetings, provided information, supplied forms, and did anything else to encourage the patient and family to engage in an informed and collaborative decision-making process with a well-informed physician."[188]

SUPPORT also provided doctors in Phase II with information on the patient's

---

180. *Id.* at 1592.
181. *Id.*
182. *Id.* at 1593.
183. *Id.* at 1592.
184. *Id.*
185. *Id.* at 1594.
186. *Id.*
187. *Id.*
188. *Id.*

prognosis. *SUPPORT* developed a prognostic model for predicting survival rates and outcomes for the types of patients involved in the study. For 94% of the patients in Phase II, physicians received at least one written prognosis evaluation, and this written evaluation was placed in the patient's medical record in 80% of the cases.[189]

A. The Results of SUPPORT

The investigators reported on the results of Phase II: "In phase II of SUPPORT, improved information, enhanced conversation, and an explicit effort to encourage use of outcome data and preferences in decision making were completely ineffectual, despite the fact that the study had enough power to detect small effects."[190] In describing the detail that supported their conclusion, *SUPPORT* investigators reported that only 34% of the physicians acknowledged having received a report of patient preferences, although a written report was provided in 78% of the cases.[191] Fifty-nine percent of the doctors acknowledged having received the prognosis report, even though a written report was given to the doctor in 94% of the cases and had been placed in the patient's medical record in 80% of the cases.[192] Only 15% of the participating doctors reported having discussed the information on prognosis and preferences with their patients.[193] The prevalence and timing of written DNR orders was the same for the control and the intervention groups. The investigators reported that there was "a small association" of the intervention with improvement in patient-physician agreement on the patient's desires concerning resuscitation.[194] There was no change in the use of hospital resources or in the length of stay in the intensive care unit. There was an increase in reported untreated pain in the interviews conducted with patients themselves.[195] The intervention did not change the unadjusted proportion of patients or surrogates who reported having a discussion about CPR (37% of the control group and 40% of the intervention group), with 41% of those reporting that CPR had not been discussed stating that they would have liked to discuss the decision.[196]

B. Reactions to SUPPORT

The President of the Robert Wood Johnson Foundation said that "the investigators were stunned" with the results of Phase II.[197] He reported that he was not surprised with the results, however, because of his own experiences as

---

189. *Id.* at 1595.
190. *Id.* at 1596.
191. *Id.* at 1595.
192. *Id.*
193. *Id.*
194. *Id.* at 1596.
195. *Id.*
196. *Id.*
197. Schroeder, *supra* note 170.

a physician and because of the "horror stories" told him by relatives and friends. The investigators themselves stated: "We are left with a troubling situation. The picture we describe of the care of seriously ill or dying persons is not attractive." They also described the ideal to which they compared the Phase II results: "One would certainly prefer to envision that, when confronted with life-threatening illness, the patient and family would be included in discussions, realistic estimates of outcome would be valued, pain would be treated, and dying would not be prolonged."[198] This was the investigators' vision of good care for the dying. But the investigators observed that "most patients and families indicated they were satisfied [with the care], no matter what happened to them."[199]

The investigators offered their own advice for the appropriate response to the results of *SUPPORT*. They conclude their report by arguing that improved care for the dying and improved communication require "reexamination of our individual and collective commitment to these goals, more creative efforts at shaping the treatment process, and perhaps, more proactive and forceful attempts at change."[200]

The Robert Wood Johnson Foundation, which funded *SUPPORT*, took a "proactive" and "forceful" action in reaction to the results of *SUPPORT*. The Foundation established a national campaign to improve care of the dying. This "Last Acts" campaign is a nationwide multi-million-dollar effort to improve care of the dying through funded research, demonstration projects, consensus conferences, task forces, and media events. The Foundation also funded the first round of expert commentary on *SUPPORT*, published in the Hastings Center Report to coincide with the release of the results of Phase II.[201]

C. Does SUPPORT *Answer Our Questions?*

SUPPORT has revealed a gap between behavior and normative expectations. Current normative expectations require that physicians talk with their patients or the patients' surrogates, about prognosis, treatment options and choices. The same norms expect that patients will participate in such discussions and perhaps will want to have such discussions. Norms requiring discussion include an expectation that physicians will account for, if not follow, the patients' treatment preferences. These are not new ethical and legal standards. They have provided the dominant ethical and legal framework for life-sustaining treatment decisions for the past two decades and for informed consent to treatment since at least the mid-1950s, if not earlier.[202]

---

198. *SUPPORT*, *supra* note 25, at 1597.

199. *Id.* at 1596.

200. *Id.* at 1597.

201. Schroeder, *supra* note 170.

202. Justice Cardozo's opinion in *Schloendorff v. Society of New York Hospital*, 105 N.E. 92 (1914), is often cited as the root of the right to refuse treatment. In that opinion, Cardozo stated: "Every human being of adult years and sound mind has a right to determine what shall be done with his own body." *Id.* at 129. For one history and a critique of legal enforcement of informed consent,

Does *SUPPORT* tell us that patient choice concerning life-sustaining medical treatment is an ideal that should be abandoned? *SUPPORT* certainly reveals that the intervention used in Phase II to increase conversation and information did not improve the incidence of patient-physician communication substantially; that it only minimally improved the accuracy of physician knowledge of patient preferences; that it did not alter physician decisions concerning writing a DNR order or discharging a patient from an ICU; that it did not improve treatment for pain; and that the great majority of both physicians and surviving family members reported feeling satisfied that things went well. But many of those evaluating the *SUPPORT* results charge this failure to defects in the intervention itself: the intervention put an intermediary between the physician and the patient;[203] the intervention used nurses instead of physicians as the stimulus for change;[204] the intervention did not try to alter the organizational environment and incentives for maintaining current behaviors;[205] the Project's prognosis model produced results no more accurate than physician judgment and so could not change decisions that would have been made by physicians without the model;[206] and so on.

Some argue that another intervention designed differently could succeed[207] and that the lesson of *SUPPORT* is that we must redouble our efforts to assure that behavior conforms to our ideals. In this view, *SUPPORT* is not the death knell for patient autonomy and informed consent in decisions concerning life-sustaining treatment; it is argued that more education and more commitment can change the patterns revealed by *SUPPORT*.

The more radical responses to *SUPPORT* have argued that *SUPPORT* is not groundbreaking. *SUPPORT* joins a great body of evidence and analysis that proves that informed consent, conversation and patient autonomy is ill-suited to decision making regarding medical interventions. Physicians simply will not talk with their patients, perhaps especially in terminal care, and will not yield control to the patients. Doctors may not be the only ones resistant to planning for death.[208] Some argue that it is time to abandon the myth of individual patient

---

see KATZ, *supra* note 138. *See also* Schuck, *supra* note 4, at 900 (observing that the legal doctrine did not emerge in "relatively robust form" until 1957).

203. Ellen H. Moskowitz & James L. Nelson, *The Best Laid Plans*, 25 HASTINGS CTR. REP. S3 (1995).

204. *Id.*

205. Mildred Z. Solomon, *The Enormity of the Task: SUPPORT and Changing Practice*, 25 HASTINGS CTR. REP. S28 (1995).

206. On the reliability of the prognosis report as compared with the individual physicians' judgments, see Joanne Lynn et al., *Accurate Prognostications of Death: Opportunities and Challenges for Clinicians Caring for Patients at the End of Life*, 163 W.J. MED. 250 (1995); Bernard Lo, *Improving Care Near the End of Life: Why Is It So Hard?*, 274 JAMA 1634 (1995).

207. Bernard Lo, *End-of-Life Care After Termination of SUPPORT*, 25 HASTINGS CTR. REP. S6 (1995); Solomon, *supra* note 205.

208. Ellen Goodman, *The Conspiracy of Silence About Death is Too Strong*, BALTIMORE SUN, Dec. 1, 1995, at 27A (reporting on family and patient resistance); *Special Report: Critical*

autonomy, or consumer choice, and redesign the relationship between physician and patient along different and yet to be announced lines.[209]

CONCLUSION

While researchers design the "next *SUPPORT*" to further test the workability of patient choice in end-of-life care, legislatures and courts are being asked to regulate the area. Empirical research can certainly aid in this effort, but will not resolve the ultimate issues. Empirical research does not have the answers. It is likely that the studies will never be well-enough designed or the results conclusive enough to predict the direction and impact of change.

Furthermore, the tension between what is now and what should be cannot be resolved by the data. The best that can be hoped for is that assumptions are recognized for what they are rather than accepted with the confidence of fact and where empirical data do indicate patterns of behavior, those patterns are recognized rather than denied.

What we do not know we can make up, but we should all realize when that is being done. What we do know, as little as that may be, we cannot ignore. It would be a great loss, if advocacy for legalization of physician aid in dying or assisted suicide triggered the collapse of the social pact that allows us to stop aggressive treatment of terminally ill patients, vegetative patients, or patients for whom treatment causes pain without benefit. But it also seems foolhardy to legalize assisted death on the basis of a fantasy of physician commitment to patient choice in health care and to the provision of adequate pain relief.

Times for Critical Care: What About "Quality of Death," MED. OUTCOMES & GUIDELINES ALERT, Nov. 30, 1995 (describing need for changing institutional culture rather than physician attitudes).

209. *See, e.g.*, Eric G. Anderson, *Some New Lessons to Consider About Life—and Death*, AM. MED. NEWS, Aug. 5, 1996, at 35.

JUDICIAL OPINIONS INVOLVING HEALTH INSURANCE COVERAGE: TROMPE L'OEIL OR WINDOW ON THE WORLD?

WILLIAM M. SAGE*

INTRODUCTION

In contrast to a few areas of health care law with strong traditions of research, such as antitrust and medical malpractice,[1] the contractual relationship between health insurer and insured has remained relatively untested empirically. There has been considerable research on access to health insurance and its cost, but relatively little on how insurance operates for those who have it. Studies which have been performed tend to focus on one important subset of coverage issues—disputes regarding the "medical necessity" of treatment or its "experimental" or "investigational" character—and apply an even narrower method: explaining the legal system by examining the written opinions of courts in cases they have decided.

A recent example of this genre is an ambitious, methodologically sophisticated, two-year investigation headed by Mark A. Hall and Gerard F. Anderson, and funded by the Federal Agency for Health Care Policy and Research (the "Hall study").[2] The Hall study was designed to test several hypotheses regarding judicial treatment of coverage decisions.[3] These included the effect on judicial outcomes of (i) the method of technology assessment employed by insurers, (ii) the severity of the patient-plaintiff's illness, (iii) the contractual language used in the policy, (iv) the presence of procedural protections, (v) the substantive and procedural barriers to recovery under the Employee Retirement Income Security Act of 1974 (ERISA),[4] and (vi) the changing perceptions of cost constraints in the health care system.

The authors of the Hall study performed multivariate analysis of data derived from 203 published opinions between 1960 and 1994.[5] They found that the following factors were significantly associated ($p<.05$)[6] with patients prevailing in coverage disputes: not being in federal appeals court, the insurance contract not expressly reserving interpretive discretion to the insurer, and seeking

---

* Associate Professor of Law, Columbia University. The author thanks Max Mehlman, Karen Jordan, Mark Hall and Karen Rotschafer Sage for comments and advice, and David Anderson for research assistance.

1. *See* Stephen Zuckerman et al., *Information on Malpractice: A Review of Empirical Research on Major Policy Issues*, LAW & CONTEMP. PROBS., Spring 1986, at 85.

2. *See* Mark A. Hall et al., *Judicial Protection of Managed Care Consumers: An Empirical Study of Insurance Coverage Disputes*, 26 SETON HALL L. REV. 1055 (1996).

3. Grant proposal on file with author.

4. 29 U.S.C. §§ 1001-1461 (1994 & Supp. I 1995).

5. Hall et al., *supra* note 2, at 1058.

6. A p-value of under .05 means that, statistically speaking, there is less than a one-in-twenty likelihood that the observed association was due merely to chance.

treatment for a life-threatening clinical condition.[7] The authors were themselves surprised to report that the presence or absence of ERISA was not a significant independent predictor of outcome.[8]

The crucial caveat for this type of research is less whether it yields valid answers but whether it asks the right questions. Judicial decisions are visible, but are they important? That is, do they indicate how well our largely private, pluralistic system of health insurance accomplishes its public policy goals? If not, research on judicial decisions risks the absurdity of looking for the lost coin under the lamppost solely because the light is better.

There is also an inherent paradox about empirical research involving outcomes of litigation. By the act of engaging in it, we are perhaps admitting to ourselves that the legal system is dysfunctional.[9] Research on decided cases asks why some parties "win" and others "lose." This is important information not only to actual or potential litigants, but to policymakers concerned with the efficiency of the courts as a forum for resolving disputes.

In our common law system, however, the results of litigated cases are supposed to create or confirm law. If coverage decisions generated clear, binding precedent, there would be little to analyze statistically (and few valid objections to the results of such analysis), since the cases would state the law. An important realization is, therefore, that we only "study" decisions empirically *because* the cases do not state the law.[10] In other words, empirical research on judicial decisions looks for subtexts where the text is unreadable.

This essay offers a few thoughts about using judicial decisions as the dataset for research into health insurance coverage. Part I offers a general overview of insurance coverage law. Part II considers why students of health insurance coverage gravitate toward studying published opinions. Part III then discusses what is wrong with the approach, and suggests alternatives. Finally, Part IV turns to what may be *right* with the approach, concluding that judicial opinions in coverage litigation may reveal the functionality (or dysfunctionality) of the coverage process in managed care. Although the basic critique which the essay presents applies to areas other than litigation involving medical necessity or experimental treatment, it offers special insights into issues like health insurance coverage where legal doctrine and public policy may not be congruent.

---

7. Hall et al., *supra* note 2, at 1067.

8. *Id.*

9. As discussed below, this is one reason why medical malpractice has been an intuitively appropriate subject for empirical research.

10. The grant proposal for the Hall study, *see supra* note 2, anticipated that most of the decisions considered would be from appellate courts, and concluded that this would be advantageous because such decisions were more likely to create binding legal precedent. At the same time, the study's application of statistical methods of analysis to opinions implicitly recognized that little binding precedent was being created.

I. WHY COVERAGE LAW MATTERS

Insurance coverage, long a backwater of health law, has come to the forefront in recent years for very good reasons.[11] First, more medical treatments are available, many of which may be exceedingly expensive. Marginally beneficial treatments are now frequently recommended by mainstream members of the academic medical community, where they once were the province of fringe practitioners or outright charlatans. As a result, disputes are less often about naturopathy or Laetrile, and more often about chemotherapy, immunotherapy, bone marrow reconstitution and organ transplantation for life-threatening illness.

Second, and relatedly, cost has become of significantly greater concern to sponsors of health insurance, notably employers and government. Not only has this made insurers more likely to challenge proposed therapies, but it has led to the development of a variety of prospective methods to control expenditures—techniques which fall under the general rubric of "managed care."

Finally, the nature of insurance in health care is ambiguous.[12] On one hand, it can be viewed as the efficient diversification of unsystematic but similar risks. On the other hand, it can be seen as a process of social pooling, and hence redistribution. Regulatory interventions in health care sometimes follow the former paradigm, sometimes the latter.

Coverage litigation has therefore become one of the American health system's Crimeas: a designated battleground for opposing values. On one side are arrayed individual patients with idiosyncratic needs, and the physicians and hospitals who stand ready to serve them. On the other side can be found employers, insurers and government—in each case claiming to represent the interests of beneficiaries or taxpayers as a whole by denying relief to one member of the group. This is, of course, the core challenge of managed care: creating an efficient system of population-based health management which nonetheless accounts equitably for the interests of individuals.

Litigation resulting from opportunistic behavior by insurers and 20-20 hindsight by beneficiaries is not unique to health care, but affects the insurance industry broadly. Unsurprisingly, courts considering individual controversies arising under blanket policies have occasionally strayed from clear doctrine. Professor Jeffrey Stempel lists the common elements of disputes over insurance coverage language which "tend to bring results less doctrinaire and consistent" than in other areas of law: standard form contracting, unequal bargaining power, non-negotiated terms, ambiguity, and recurring equitable considerations.[13]

---

11. *See generally* Mark A. Hall & Gerard F. Anderson, *Health Insurers' Assessment of Medical Necessity*, 140 U. PA. L. REV. 1637 (1992).

12. Deborah A. Stone, *The Struggle for the Soul of Health Insurance*, *in* THE POLITICS OF HEALTH CARE REFORM: LESSONS FROM THE PAST, PROSPECTS FOR THE FUTURE 26 (James A. Morone & Gary S. Belkin eds., Duke Univ. Press 1994).

13. JEFFREY W. STEMPEL, INTERPRETATION OF INSURANCE CONTRACTS: LAW AND STRATEGY FOR INSURERS AND POLICYHOLDERS § 2.1 (Brown & Little 1994) (quoted in Peter Nash Swisher, *Judicial Interpretations of Insurance Contract Disputes: Toward a Realistic Middle*

Managed care has intensified the problems facing courts asked to determine coverage. These derive principally from three sources: the fact that health insurance in the United States is governed by disparate bodies of law, the convergence of coverage and care in prepaid systems, and the increasing risk of conflicts of interest affecting payers and providers.

A. Same Problem, Different Rights

Well-insured patients with identical medical conditions seeking equivalent treatments are far from equals in a court of law. The relatively small group of individual policyholders and employees of state and local governments with private insurance are protected by state insurance regulations and have available a panoply of legal claims and remedies under state law if coverage disputes arise. The much larger group of persons insured through their workplace under ERISA have a limited set of rights and remedies, and the subset whose employers self-insure are denied the benefit of state regulatory intervention as well.[14] Beneficiaries of government programs such as Medicare and Medicaid are subject to procedural restrictions on judicial review in addition to slightly different substantive standards.[15] Federal workers receiving coverage through the Federal Employees Health Benefit Program (FEHBP),[16] active duty military under the Civilian Health and Medical Program of Uniformed Services (CHAMPUS),[17] and veterans covered by the Veterans' Administration health system are also treated somewhat differently than privately insured individuals.[18]

Moreover, these financing systems increasingly deliver services through the same corporate managed care entities. Consequently, an individually insured person, a worker whose employer purchases insurance, an employee of a self-insured firm, and a retiree who has enrolled in the Medicare product offered by her insurer may think they are dealing with the same "health plan." In fact, they may be subject to widely disparate rules. Because managed care organizations operate under readily apparent cost constraints, this different legal treatment of similarly situated individuals tends to offend notions of basic fairness.

B. Convergence of Coverage and Care

Coverage litigation prior to managed care was perceived as tangential to health care delivery. Disputes generally arose long after treatment had been rendered, and focused on payment rather than survival. Not only did this allow

---

Ground Approach, 57 OHIO ST. L.J. 543, 560 (1996)).

14. Pilot Life Ins. Co. v. Dedeaux, 481 U.S. 41 (1987); *see also* Margaret G. Farrell, *ERISA Preemption and Regulation of Managed Health Care: The Case for Managed Federalism*, 23 AM. J.L. & MED. 251 (1997).

15. *See* Eleanor D. Kinney, *The Medicare Appeals System for Coverage and Payment Disputes: Achieving Fairness in a Time of Constraint*, 1 ADMIN. L.J. 1 (1987).

16. 6 U.S.C. §§ 8901-8914 (1994 & Supp. I 1995).

17. 10 U.S.C. §§ 1071-1106 (1994).

18. *See* 44 AM. JUR. 2D *Insurance* § 2057 (1982).

reviewing courts some detachment from compelling human equities, but it meant that health care policymakers did not connect coverage litigation with broader themes of access, cost and quality.

As the cost of medical care has risen, and as managed care organizations have combined financing with delivery of services, the relationship between coverage denials and inability to receive care has become clearer. In particular, precertification requirements for hospitalization and surgery have increased the urgency of resolving disputes, as well as raising the stakes for plaintiffs. Similar considerations apply when patients seek access to specialty services not readily available within a managed care network. At the same time, insurance organizations with selective physician networks, drug formularies, and strict utilization review standards are more likely to be implicated in patient injury resulting from coverage denied or improperly provided.[19]

C. Conflicts of Interest

By incorporating financial incentives for cost containment into provider contracts or otherwise promoting physician compliance with organizational goals, managed care has realigned the traditional parties to coverage disputes. For example, the recommendations of the physicians affiliated with the health plan and other providers consulted by the patient may differ. At the extreme, patients in managed care systems not only may be unsure of their insurer's financial obligations, they may be unaware that potentially beneficial treatment exists.

Informational asymmetries between managed care organizations and individual patients relating to coverage have taken on correspondingly greater importance. Unethical marketing practices and other potential opportunities for insurers to deceive beneficiaries have always influenced courts. Conventional insurance, however, presented limited opportunities for mischief. By contrast, coverage law in the era of managed care must monitor a broad array of intermediaries, including not only traditional insurance agents but also employee benefits personnel, claims administrators and utilization review entities, plus affiliated providers and their contracting vehicles.

Taking these factors in combination, coverage disputes are increasingly about providing fair and uniform access to medical treatment from competent agents making full disclosure. The world of insurance coverage law thus begins to resemble the more familiar arena of medical malpractice. In the discussion that follows, we will come back to this comparison in connection with the pros and cons of, and alternatives to, studying coverage law through reported cases.

II. WHY WE STUDY COVERAGE LAW USING REPORTED DECISIONS

At first glance, studying coverage law from judicial decisions strikes one as

---

19. *See, e.g.*, Wilson v. Blue Cross, 271 Cal. Rptr. 876 (Cal. Ct. App. 1990). *But see* Corcoran v. United Healthcare, Inc., 965 F.2d 1321, 1331 (5th Cir. 1992) (ERISA preempts available remedies for "medical decisions incident to benefit determinations").

sensible. Contractual provisions designed to apply across-the-board seem susceptible to legal interpretation apart from the facts of particular cases. For example, in the case of clauses excluding coverage of "investigational" or "experimental" treatment, the therapy, not the patient, is the apparent focus of inquiry. Therefore, judicial interpretation of insurance provisions could have systematic effects.[20]

These expectations are not borne out in practice. Despite the fact that insurance is an aggregate endeavor, and the policy language interpreted by courts affects many people simultaneously, reported cases generally reflect unique needs and circumstances. As a result, coverage cases are seldom brought in class action form. Moreover, although a ruling overturning an exclusion for experimental treatment could theoretically change the policy for all patients requesting the same treatment, this does not seem to occur.[21] Why, then, are reported opinions an attractive database for empirical research?

A. Judicial Decisions are Abundant

A major reason we tend to examine judicial opinions in coverage cases is that there seem to be a lot of them.[22] However, the apparent abundance of formal judicial findings is largely artifactual.

Most personal injury cases are tried before juries.[23] However, compared with medical malpractice disputes, there are relatively few jury verdicts in coverage disputes, a phenomenon probably explainable by ERISA. Though the percentage has declined in recent years, well over half of privately insured patients receive coverage through employment.[24] Coverage claims involving employer-sponsored

---

20. By contrast, medical malpractice cases tend to depend on highly idiosyncratic facts and, as discussed further below, are generally tried before juries whose reasoning is not revealed in written opinions. Some empirical work in malpractice has therefore explored the factors that cause plaintiffs to win or lose in jury verdicts and settlements. Moreover, many empirical studies of malpractice have focused on non-judicial measures of system performance, assessing global cost and efficiency from data such as malpractice insurance premiums, defensive medicine, correlation between negligent injury and litigation, and adequacy of compensation. *See infra* note 114.

21. In addition, a defendant insurer could be collaterally estopped from challenging a prior adverse determination regarding the meaning or legitimacy of a contractual exclusion. *See generally* Parklane Hosiery Co. v. Shore, 439 U.S. 322 (1979). *See also* Clements v. Airport Auth., 69 F.3d 321, 330 (9th Cir. 1995); Texas Employers' Ins. Ass'n v. Jackson, 862 F.2d 491, 500 (5th Cir. 1988); 18 JAMES WILLIAM MOORE, MOORE'S FEDERAL PRACTICE § 132.01 (3d ed. 1997). Nonetheless, courts are reluctant to extend rulings from one case to another except through the more limited application of stare decisis. Consequently, although insurers may redraft contractual provisions in response to judicial decisions, they are seldom compelled to.

22. This seeming abundance may be deceptive when samples are subjected to rigorous statistical analysis. *See infra* Part III.A.

23. Kevin M. Clermont & Theodore Eisenberg, *Trial by Jury or Judge: Transcending Empiricism*, 77 CORNELL L. REV. 1124, 1136-37 (1992).

24. Mark A. Hall, *Rationing Health Care at the Bedside*, 69 N.Y.U. L. REV. 693, 779

health plans must be brought under federal ERISA law, which broadly preempts related state law claims.[25] Although the Seventh Amendment to the U.S. Constitution guarantees the right to a federal jury trial upon request of a party in "suits at common law,"[26] disputes involving equitable remedies may be conducted as bench trials.[27] Claims for benefits under ERISA were traditionally regarded as equitable because federal pension law incorporates large portions of the law of trusts.[28]

As a result, several federal appellate courts have denied the right to a jury trial in ERISA cases, although there are indications this may change with respect to claims for damages as opposed to injunctive relief.[29] Bench trials obligate the judge to issue a Memorandum of Findings of Fact and Opinions of Law, which may be published.[30] By contrast, a jury verdict in a medical malpractice case creates little official record, unless detailed rulings are issued on post-trial motions.

Whether or not brought under ERISA, suits in preauthorization cases requesting preliminary injunctions requiring insurers to pay for or provide treatment are decided by judges. In addition, a subset of coverage disputes arises under Medicare and Medicaid, both historically fee-for-service insurance programs. Challenges to Medicare or Medicaid benefit determinations are channeled through administrative adjudicatory mechanisms.[31] Because beneficiaries have a right to judicial review (although not to a jury trial), many of these cases result in published opinions. Moreover, settlement opportunities

---

(1994).

25. 29 U.S.C. §§ 1144(a) & 1003(a) (1994).

26. U.S. CONST. amend. VII.

27. *See* Katchen v. Landy, 382 U.S. 323, 337 (1966) (citing Barton v. Barbour, 104 U.S. 126, 133-34 (1891)).

28. *See* Coar v. Kazimar, 990 F.2d 1413, 1418 (3d Cir. 1993).

29. *See, e.g.*, Katsaros v. Cody, 744 F.2d 270 (2d Cir. 1984). However, the Supreme Court recently suggested that the nature of the desired remedy, not the underlying claim, may determine the right to a jury trial. Mertens v. Hewitt Assoc., 508 U.S. 248 (1993) (denying claim for damages as not within category of "appropriate equitable relief"). Following *Mertens*, several federal district courts have required jury trials for benefits claims requesting damages. *See, e.g.*, Mullins v. Pfizer, Inc., 889 F. Supp. 69, 76 (D. Conn. 1995); Hulcher v. United Behavioral Sys., Inc., 919 F. Supp. 879, 885 (E.D. Va. 1995); Algie v. RCA Global Communications, Inc., 891 F. Supp. 870, 875 (S.D.N.Y. 1994); Sullivan v. LTV Aerospace & Defense Co., 850 F. Supp. 202 (W.D.N.Y. 1994), *vacated in part by* 82 F.3d 1251 (2d Cir. 1996).

30. According to the *Bench Book for U.S. District Court Judges*, which sets forth guidelines for when trial courts are required to issue Findings of Fact and Conclusions of Law in civil cases and motions, bench trials must result in written findings, as must granting or refusing interlocutory injunctions (e.g., preliminary injunctions requiring coverage). A written opinion generally must also accompany a grant or denial of summary judgment. FEDERAL JUDICIAL CENTER, BENCH BOOK FOR U.S. DISTRICT COURT JUDGES, § 6.02 (4th ed. 1996).

31. *See* Eleanor D. Kinney, *Procedural Protections for Patients in Capitated Health Plans*, 22 AM. J.L. & MED. 301, 314-18 (1996).

are rare in public programs.

Motion practice also exaggerates the frequency of judicial intervention in coverage cases. Even if a jury might theoretically be impanelled for an ERISA benefits trial, courts are limited to an abuse of discretion review in situations in which the ERISA plan document expressly reserves discretion to the plan administrator.[32] Defendants therefore typically bring motions for summary judgment, which often generate written rulings. Additionally, some jury decisions in non-ERISA cases brought under state law may result in the assessment of punitive damages (which are rare in malpractice cases), occasioning post-trial motions to remit damages which must be decided as a matter of law by the trial judge.[33]

Apart from identifying parties as plaintiffs, defendants, appellants or appellees, the Hall study does not detail the procedural posture of the cases it reviewed. Nonetheless, the distribution of cases between trial and appellate levels in the Hall study is suggestive. State supreme court decisions accounted for 9% of the opinions reviewed, state appeals courts 26%, state trial courts 3%, federal appeals courts 22% and federal trial courts 39%.[34] The larger percentage of federal trial court decisions may represent not only Medicare and Medicaid, but bench trials or motions in ERISA cases. The low percentage of state trial court opinions, on the other hand, may reflect non-ERISA cases decided by state juries. Assuming that juries favor plaintiffs, this hypothesis is supported by the fact that the insurer was the appellee in most federal appeals (66%) but was the appellant in most state appeals (55%).[35]

B. Contract Cases Appear Self-Contained

The contractual nature of coverage disputes may favor research using judicial decisions. Parties to a contract form a voluntary relationship, the terms of which are subject to judicial enforcement but little more. Compared with tort claims, which convey intuitively a need to study broader social issues such as deterrence of negligence and compensation for injury, contract cases seem less concerned with factors beyond the agreement. In other words, tort analysis may predispose to extrinsic research because public policy issues are explicit. By contrast, such matters are implicit in contract analysis, which therefore tends to limit research to intrinsic data such as judicial interpretation.

This is not to say that coverage language is written on a blank slate. Insurance benefits provided by fee-for-service Medicare and Medicaid are not

---

32. Firestone Tire & Rubber Co. v. Bruch, 489 U.S. 101, 115 (1989).

33. There is another possible reason why coverage litigation often results in a final judicial determination. Lawsuits in coverage cases frequently represent last-ditch efforts by dying patients and their families. Unlike malpractice cases, which center on monetary compensation for prior injury, managed care litigation is all-or-nothing. The potential for compromise, and therefore for settlement, may therefore be reduced.

34. Hall et al., *supra* note 2, at 1061.

35. *Id.* at 1064.

contracts but legislative entitlements. In addition, the terms of coverage are frequently constrained by state insurance regulation.[36] Nonetheless, strong ERISA preemption of state regulation for self-insured employers, coupled with the absence of federal substantive requirements, has created a period of contractual free rein for many managed care agreements. Taking advantage of this legislative laissez faire, managed care plans and employers have greatly increased the range and significance of contractual limitations imposed on beneficiaries.[37]

As managed care becomes the dominant form of insurance, and coverage and care converge, the current era of free contracting will probably come to an end.[38] Most importantly, federal legislators are beginning to amend ERISA to impose on managed care plans substantive limitations typical of state insurance regulation.[39] At the same time, both federal and state regulators seem more willing to dictate medical practice in the context of managed care than was the case in a fee-for-service environment.[40]

C. Academics Understand Judicial Opinions

Quantitative and statistical work in law is a relatively recent phenomenon,

---

36. THE POLITICS OF HEALTHCARE REFORM: LESSONS FROM THE PAST, PROSPECTS FOR THE FUTURE 51, 210 (James A. Morone & Gary S. Belkin eds., 1994). Insurance regulation includes limitations on contractual language as well as issues such as reserve requirements and mandated benefits. This has two consequences for judges evaluating insurance contracts. First, it reduces the range of possibilities available to contracting parties. In addition, however, it may create situations where compliance with regulatory requirements contradicts judicial principles of contract interpretation. For example, a court might berate an insurer for failing to state explicitly that a particular treatment was excluded from coverage, despite the fact that "laundry list" exclusions are disfavored by state regulators.

37. *See, e.g.*, McGann v. H & H Music Co., 946 F.2d 401, 407-08 (5th Cir. 1991) (allowing employer to adopt self-insured ERISA plan with greatly reduced benefits for AIDS).

38. One aspect of the convergence of coverage and care tending in the opposite direction is the possibility that the professional practice standard to which physicians are held might be specified in a managed care contract. In *Dukes v. U.S. Healthcare, Inc.*, for example, the court left open the possibility that a contractual standard of care would allow insurers to claim ERISA preemption even in routine malpractice litigation. 57 F.3d 350 (3d. Cir.), *cert. denied*, 116 S. Ct. 564 (1995). It will be interesting to see if contractual standards of care persuade researchers on medical malpractice to pay *more* attention to judicial opinions.

39. *E.g.*, Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (codified in scattered sections of 18 U.S.C., 26 U.S.C., 29 U.S.C., and 42 U.S.C.); *see also* Patient Access to Responsible Care Act of 1997, H.R. 1415, 105th Cong., 1st Sess.

40. *E.g.*, Newborns' and Mothers' Health Protection Act of 1996, Pub. L. No. 104-204, 110 Stat. 2874, 2935-44 (codified at 29 U.S.C. § 1185, 42 U.S.C. §§ 300gg-4, 300gg-51 (1994)) (minimum hospital length of stay for postpartum care); Breast Cancer Patient Protection Act of 1997, H.R. 135, 105th Cong., 1st Sess. (minimum hospital stay for mastectomy and lymph node dissection).

reflecting lawyers' increasing level of engagement with the world outside the courtroom. "Law and empiricism" follows naturally from previous scholarly movements—such as legal realism, law and society, law and economics, and critical legal studies—that drew upon extrinsic sources of information and analysis to explain and inform legal doctrine. The enormous expansion of the American health care industry, and its high degree of regulation, make empirical work in health law especially attractive.

Judicial decisions are a natural starting point for legal empiricism. Judge Posner finds it significant that legal scholars focus on the written opinion rather than the courtroom drama: "[Academics tend] to ascribe more importance to the opinion, to its reasoning, its rhetoric, etc. than to the decision itself. Yet these are secondary factors for most judges. For the judge, as for Hamlet, `the play's the thing.'"[41] Moreover, Posner notes that legal academics tend to study appellate decisions more intensively than those of trial courts, and that "opinions are virtually [the] only public product" of appellate judges.[42] Mixing trial and appellate cases in empirical analyses therefore raises questions. According to Posner, the two tiers of judging differ significantly: trial judges are both affected and monitored by daily interaction with the litigants, while appellate judges play a "game" according to intellectually satisfying but more formalistic rules.[43]

Published decisions therefore represent an easy extension of traditional legal scholarship to health care. With respect to coverage litigation specifically, ERISA's broad preemptive effect has probably encouraged this focus by reducing the number and influence of legislative initiatives. This is likely to change as the "ERISA vacuum" begins to fill in response to recent judicial limits on preemption and heightened interest in federal regulation of insurance.[44]

D. Additional Data Are Limited

Availability of information drives the direction of research. Reported cases are the most easily available source of information about coverage disputes, merely an electronic search away. Some data on a broader section of coverage disputes are also available, at least for federal cases. The Administrative Office of the United States Courts creates a record of each civil case terminated, including the subject matter and jurisdictional basis, the amount demanded, the dates of filing and termination, the procedural posture of the case at termination and, if a judgment was reached, the prevailing party and the amount awarded.[45]

---

41. Richard A. Posner, *What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does)*, 3 SUP. CT. ECON. REV. 1, 26 (1993).

42. *Id.* at 7.

43. *Id.* at 7, 29-30.

44. *E.g.*, Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (codified in scattered sections of 18 U.S.C., 26 U.S.C., 29 U.S.C., and 42 U.S.C.) (also known as the Kassebaum-Kennedy Act; applies to self-funded ERISA plans as well as insurance arrangements).

45. Clermont & Eisenberg, *supra* note 23, at 1133.

Whereas lawsuits are matters of public record, relatively little is known about how insurers reach coverage decisions in non-litigated cases.[46] Unlike medical malpractice, where many liability insurers freely share data on their insureds, health insurers often assert proprietary interests in their coverage standards.[47] Even coverage determinations under government programs are obscure, largely a result of the low visibility of the private carriers and intermediaries with which Medicare contracts.[48] This may change if constitutional due process requirements are imposed on managed care organizations serving Medicare and Medicaid beneficiaries.[49]

By contrast, medical malpractice data are far more organized and accessible. Research on malpractice tracks malpractice insurance premiums, claims and awards. For example, data on nationwide jury awards are available from Jury Verdict Research, although the database depends on voluntary submissions by the parties and excludes settlements.[50] Other important sources of information are malpractice insurers' records of filed and closed claims, and state and federally mandated reporting of settlements and judgments involving physicians.

The new regulatory focus on managed care should expand data availability in the coverage arena. For example, several states now require managed care organizations to disclose to beneficiaries information on grievance and appeals procedures for challenging coverage determinations.[51] In addition, a few states require standardized reporting to regulators and the public of patient satisfaction with coverage, including disenrollment statistics and the number and outcome of benefit disputes. This information, however, is not necessarily indexed to filed or decided litigation.

E. Political Constituencies Are Underdeveloped

Politics explains some of these data differences. Medical malpractice has

---

46. *See, e.g.*, William P. Peters & Mark C. Rogers, *Variation in Approval by Insurance Companies of Coverage for Autologous Bone Marrow Transplantation for Breast Cancer*, 330 NEW ENG. J. MED. 473 (1994). One exception is technology assessment, which has generated a rich literature. *See, e.g.*, INSTITUTE OF MEDICINE, COMMITTEE FOR EVALUATING MEDICAL TECHNOLOGIES IN CLINICAL USE, ASSESSING MEDICAL TECHNOLOGIES (1985). However, the science of technology assessment tends to be divorced from its utility in avoiding or resolving coverage disputes.

47. Insurers may also have an interest in preserving ignorance, since the alternatives might be to develop more rational standards or to admit that the emperor lacks clothes.

48. *Medicare Part B: Inconsistent Denial Rates for Medical Necessity Across Six Carriers: Before the Subcommittee on Regulation, Business Opportunities, and Technology, Committee on Small Business, House of Representatives, reprinted in* GAO/T-PEMD-94-17 (statement of Eleanor Chelimsky, Assistant Comptroller General).

49. *See, e.g.*, Grijalva v. Shalala, 946 F. Supp. 747 (D. Ariz. 1996).

50. Stephen Zuckerman et al., *supra* note 1, at 90.

51. *See* William M. Sage & David Anderson, *Health Care Disclosure Requirements*, *in* 1997 HEALTH LAW HANDBOOK 185, 189-90 (Alice G. Gosfield ed., Clark Boardman Callaghan 1997).

clearly demarcated, well organized constituencies on both sides, which has encouraged research to support desired legislative reforms and has also provided a political outlet for studies which have been performed.[52] Trial lawyers and consumer groups have stressed patient protection and the identification of "bad doctors," while physicians and hospitals have emphasized the aggregate cost and inefficiency of litigation.[53]

Regarding private insurance coverage, at least, political constituencies have thus far had less certain turf and less mature strategies. In particular, there is as yet no group (save employers, who are anxious to downplay their potential conflicts of interest as sponsors of health care benefits) arguing that there is "too much" coverage litigation or that the threat or outcome of it unacceptably increases health care costs. Insurers have kept a low profile, relying on ERISA and other arcane legal safeguards, rather than engaging in a debate between statistical and identified lives which they would surely lose.[54]

Although coverage research is still a character in search of an author, there are now several likely candidates. For example, three of the ten largest jury verdicts of 1995 involved managed care.[55] As a result, trial lawyers who had long ignored medical malpractice cases because of legislative caps on damages and the need for special medical expertise are looking twice at corporate managed care defendants and the potential for lucrative claims of "bad faith" insurance denial or infliction of emotional distress.[56] Academic health centers, which feel increasingly vulnerable as both explicit government funding and private cross-subsidies for clinical research are reduced, are another important constituency for coverage research, as well as a significant source of intellectual horsepower. In response, the insurance industry can be expected to commission or encourage its own studies of the cost of mandated benefits, required appeals

---

52. Though not the intention of its authors, for example, the most recent attempt to quantify defensive medicine immediately attracted the attention of tort reformers both within and outside of the medical community. Daniel Kessler & Mark McClellan, *Do Doctors Practice Defensive Medicine?*, 111 Q.J. ECON. 353 (1996). Obviously, widespread public interest is both an opportunity and a risk. Hensler notes that "[t]he highly politicized world of policy research challenges researchers to keep their political personae separate from their research analytic personae." Deborah R. Hensler, *Researching Civil Justice: Problems and Pitfalls*, LAW & CONTEMP. PROBS., Summer 1988, at 55, 65.

53. One should note that the former is deliberately anecdotal, while the latter generally reduces the emotional appeal of severe patient injury by making dollars rather than lives the mode of discourse.

54. At least in medical malpractice, the counterpoint to a severely injured victim is a single physician making an individualized judgment—not a faceless corporation dictating aggregate policies—although this may be changing now that incidents of malpractice can be gathered together under the perceived responsibility of managed care organizations.

55. *Verdict Trends in 1995 Disregard the Clamor: Congressional Tort Reform Has Had Little Effect*, NAT'L L.J., Feb. 5, 1996, at C2.

56. As in malpractice, moreover, the plaintiff's side generally prefers the sympathetic anecdote to the bigger picture—except, of course, to say that "it could happen to you."

process, or unrestrained litigation, as has been the case with more actively documented areas of insurance such as workers' compensation and automobile no-fault. Moreover, with all fifty states and the federal Congress firmly committed to consumer protection in managed care, the legislative can of worms (despite ERISA) has been opened wide. This should greatly increase interest in empirical studies of insurance contracting practices.

III. LIMITATIONS OF THE JUDICIAL DATASET

Empirical studies of judicial decisions suffer from significant limitations. Some of these pitfalls apply generally, while others take on special importance in health insurance coverage. As a result, these studies may not prove what they set out to prove—who prevails in coverage disputes and why.

A. Small Sample Size

Although there may be enough reported decisions in insurance coverage to attract attention, there are far too few to draw statistically meaningful conclusions except for very general issues. In some situations, this small numbers problems may prevent identification of trends that would reach statistical significance given a larger sample. In other cases, aggregating data to achieve statistical validity obscures important local variations, such as between courts in different jurisdictions.

In the Hall study, for example, the authors were surprised that only 203 cases relating to medical appropriateness and otherwise meeting their criteria for inclusion resulted in published federal and state court decisions between 1960 and 1994.[57] Because of their small sample size, the investigators were forced to abandon their original goal of performing a longitudinal analysis of cases to measure judicial responses to improved practices by the insurance industry or changing perceptions of the health care system.[58] Neither were they effectively able to study narrow but important questions identified in their original grant proposal such as the relationship between payer and outcome or the role of specific methods of technology assessment.[59]

B. Long Time Lags

Like observational astronomy, reported cases reveal the universe as it was, not as it is today. Courts are necessarily reactive, weighing in only in identifiable, fully developed controversies.[60] In the Hall study, the median time to final disposition was 2.5 years, and many cases took much longer.[61] In

---

57. Hall et al., *supra* note 2, at 1059.

58. Grant proposal on file with author.

59. Grant proposal on file with author.

60. SHELDON GOLDMAN & AUSTIN SARAT, AMERICAN COURT SYSTEMS: READINGS IN JUDICIAL PROCESS AND BEHAVIOR 8 (S. Goldman & A. Sarat eds., 1978).

61. Hall et al., *supra* note 2, at 1060.

addition, sample size constraints required pooling of cases that would have been "old" no matter how quickly they had been resolved.[62]

The time delays inherent in judicial decision making create special problems for studies of industries like managed care which are in rapid transition. Examining judicial decisions in coverage cases may therefore suffer from irrelevance. For example, only six of the 203 cases studied by Hall and Anderson involved HMOs or other managed care plans.[63]

Given the inevitable time lag, the lack of cohesiveness to current coverage law may even be a blessing in disguise. As discussed below, judges prefer to articulate narrow justifications for their decisions.[64] Because of their unusual facts and compelling equities, health care coverage cases often present an extreme example of this phenomenon. In managed care, changes in industry practice are occurring so rapidly that judicial attempts to make sweeping law would seldom synchronize with the state of the system at the time of the ruling.

C. Selection Bias

Just because there are a lot of coverage cases with reported opinions doesn't mean that most cases generate reported opinions, that most disputes give rise to litigation, that most coverage denials are disputed, or even that treatment options which might be denied are proposed in the first place.[65] Reported decisions are the tip of a very large iceberg. What occurs outside the purview of the courts is probably far more significant to the average patient—and therefore to public policymakers—than the opinions of judges.

There are many reasons why a litigated case might result in a final decision. It may indicate the failure of negotiation or an alternative mode of dispute resolution to achieve settlement or determine rights.[66] It may reflect uncertainty about the underlying law. It may indicate that the law, though clear, is

---

62. *Id.* at 1059-60.

63. *Id.* at 1056. The authors attributed this to another problem with collecting reported decisions, selection bias from unlitigated cases, but it undoubtedly relates as well to the novelty of many managed care arrangements.

64. *See infra* notes 86-96 and accompanying text.

65. In medical malpractice cases, for example, Danzon has examined the relationship between court outcomes and settlements. She concluded that claims tried to verdict involve atypically large dollar amounts, more uncertainty about liability, and weaker evidence for plaintiffs. PATRICIA M. DANZON, MEDICAL MALPRACTICE: THEORY, EVIDENCE, AND PUBLIC POLICY 50-51 (1985). Studies of legal decisions in other areas have also had to address sample selection problems. *See, e.g.*, Philip D. Drake & Michael R. Vetsuypens, *IPO Underpricing and Insurance Against Legal Liability*, FIN. MGMT., March 22, 1993, at 64.

66. Gross and Syverud assert that despite liability insurance for defendants and contingent fees for plaintiffs, trials are too expensive and risky for most parties. They conclude that "[t]he main function of trials is not to resolve disputes but to deter other trials." Samuel R. Gross & Kent D. Syverud, *Don't Try: Civil Jury Verdicts in a System Geared to Settlement*, 44 UCLA L. REV. 1, 63 (1996).

objectionable to one or both parties. It may mean that the underlying facts have not been elucidated to the satisfaction of the litigants. Finally, it may suggest that interest groups are actively pursuing an available avenue for legal change.

At the same time, trends in legal doctrine undoubtedly influence the way parties behave when they make private ordering decisions.[67] This is what Mnookin and Kornhauser refer to as "bargaining in the shadow of the law."[68] In addition, as Mather describes, trial courts are "cumulative policy makers," with the outcomes of earlier cases prompting or deterring additional, similar litigation.[69] This is enhanced by the form of discourse in litigation, where parties frequently argue similarities to or differences from previous cases in order to obtain the desired result. Nonetheless, it is a leap of faith to conclude that cases not yielding final opinions—most of which are not even "cases" in the technical sense—mirror those that do.

Selection bias can significantly skew research findings. In the Hall study, for example, 57% of the cases that had definitive outcomes were resolved for the plaintiff.[70] This may mean that coverage law favors patients over insurers. However, it may equally signify the opposite—that insurers are unwilling to settle a large percentage of valid claims, perhaps because of advantages such as ERISA's limitations on damages—or a range of intermediate positions.[71] The essential point is to recognize that decided cases are not necessarily representative of the universe of actual and potential disputes.

Even the apparent direction of legal change may be misleading. Henderson

---

67. In a recent General Accounting Office survey of HDC-ABMT for breast cancer, nine of twelve insurers who decided to cover the procedure reported that litigation or the threat of it was a factor in their decision, and five characterized legal concerns as among the most important reasons for coverage. U.S. GAO, HEALTH INSURANCE: COVERAGE OF AUTOLOGOUS BONE MARROW TRANSPLANTATION FOR BREAST CANCER, *microformed on* GA 1.13:HEHS 96-83, at 9 (GAO Documents).

68. Robert H. Mnookin & Lewis Kornhauser, *Bargaining in the Shadow of the Law: The Case of Divorce*, 88 YALE L.J. 950 (1979).

69. Lynn Mather, *The Fired Football Coach (Or, How Trial Courts Make Policy)*, *in* CONTEMPLATING COURTS 179 (Lee Epstein ed., 1995).

70. Hall et al., *supra* note 2, at 1062.

71. Priest and Klein argue that selection bias should result in an approximately 50-50 division among decided cases, which is open to misinterpretation as indicating unsettled law. George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 5-6 (1984). Other researchers have extended this reasoning. For example, Clermont and Eisenberg compared bench trials with jury trials, and concluded that plaintiffs in two areas, product liability and medical malpractice, prevailed at a much higher rate before judges. They hypothesized that when these types of personal injury cases come before judges, defendants are overly confident and decline settlement opportunities, so that plaintiffs win a larger percentage of ultimate judgments. Clermont & Eisenberg, *supra* note 23, at 1162. *See also* Robert H. Gertner, *Asymmetric Information, Uncertainty and Selection Bias in Litigation*, 1993 U. CHI. L. SCH. ROUNDTABLE 75 (concluding that information asymmetrics can explain deviations from the 50-50 rule).

and Eisenberg point out that a change in the percentage of cases won or lost by each side explains little; for example, plaintiffs may lose a larger fraction of decisions over time because they are bringing more cases in response to a *favorable* change in the law, while defendants are settling weaker cases more frequently and trying only the stronger ones.[72]

Another important aspect of selection bias is its susceptibility to deliberate manipulation. For example, lower-cost forms of alternative dispute resolution such as mandatory, binding arbitration may be attractive to managed care organizations. However, these methods exist at the sufferance of the legal system, since submitting to binding arbitration implies a waiver of one's right of access to the courts. Balance among outcomes is a superficial indication of impartiality. Planning the organization's settlement strategy to produce an even split in decided cases may convince a reviewing court or legislature that a biased dispute resolution process is in fact fair, and may therefore discourage it from tinkering with or overturning it.

The authors of the Hall study discuss case selection issues in connection with the limited sample size. For example, they point to both long delays in resolution and the fact that the median cost of treatment at issue was between $10,000 and $50,000 as deterrents to litigating cases,[73] especially since ERISA generally limits damages to the value of the benefit denied.[74] In addition, they speculate that managed care gives rise to fewer litigated cases because many denials take place at the treating physician or supervising physician level, reducing patients' knowledge of their options.[75]

Recall that the Hall study concluded that whether a case is governed by ERISA is not a significant predictor of outcome.[76] This finding suggests that another important selection bias may have escaped detection. Of the cases studied, 17% were Medicare, 13% Medicaid, 34% commercial insurance, 18% Blue Cross, 7% self-insured, 3% Taft-Hartley and 7% FEHBP or CHAMPUS.[77] The authors do not indicate what percentage of the commercial insurance and Blue Cross cases involved insured ERISA plans. Nonetheless, given that nearly half of employers self-insure,[78] the small number of self-insurance cases indicates

---

72. James A. Henderson, Jr. & Theodore Eisenberg, *The Quiet Revolution in Products Liability: An Empirical Study of Legal Change*, 479 UCLA L. REV. 479, 502 (1990).

73. Hall et al., *supra* note 2, at 1060.

74. 29 U.S.C. § 1132(a)(1)(B) (1994).

75. Hall et al., *supra* note 2, at 1061. The Hall study recognizes other selection biases as well. For example, the authors interpret the low win rate for patients with life-threatening conditions as evidence that insurers are more cautious about denying coverage to these patients. *Id.* at 1065.

76. On the other hand, the study found that federal appellate jurisdiction and contractually reserved discretion by the insurance plan favor defendants. *Id.* at 1067. These factors are closely linked to ERISA, making it problematic to consider them independent variables.

77. *Id.* at 1061.

78. *Managed Care: HMOs, PPOs, POs Now Cover Majority of Americans in Employer Plans*, 24 Pens. & Benefits Rep. (BNA) 316 (1997). Alternatively, the low percentage of self-

that ERISA plan beneficiaries are underrepresented in cases resulting in judicial decisions.

If most ERISA cases (or potential disputes) never reach decision, ERISA becomes a very important factor regardless of the outcome of reported opinions. As noted previously, ERISA restricts claims and damages, seldom confers a right to a jury trial, and limits judicial review in many instances.[79] Anecdotal evidence exists that many complaints are not pursued if defense counsel responds to claims as being preempted by ERISA. This strongly suggests that ERISA is a powerful deterrent to suit, and therefore a predictor of outcome as it should inform public policy.

D. Publication Bias

In addition to selection bias, judicial decisions suffer from reporting bias. Medical researchers are more likely to publish studies establishing causation or clinical benefit than ones demonstrating its absence,[80] Similarly, judges publish only a fraction of the opinions they write.[81] As a result, legal reporters and on-line databases include a preponderance of rulings containing groundbreaking legal analysis or novel conclusions. Among other things, this tendency can lead legal researchers to overestimate the mutability and drama of the law.

Publication bias can also affect the geographic distribution of cases, which can change modal conclusions regarding the law. State appellate courts vary considerably in their publishing practices. Florida, for example, publishes more than three times as many opinions as California despite its much smaller population.[82] Overall, state courts are generally more predisposed to publish their holdings than federal appeals courts, which have instituted fairly uniform controls on publication.[83]

Appellate opinions are most clearly biased in favor of novelty. For example, Rule 53 of the Seventh Circuit Court of Appeals requires published opinions when the decision (i) establishes a new, or changes an existing rule of law; (ii) involves an issue of continuing public interest; (iii) criticizes or questions

---

insured cases may reflect dilution of the data set by cases from earlier decades when self-insurance was rare.

79. *See supra* notes 14, 23-29, and accompanying text.

80. Publication bias is widely recognized in medical research, where the intellectual (and sometimes financial) appeal of affirmative results leads to a high frequency of false positives in the clinical literature. *See* Kay Dickersin, *The Existence of Publication Bias and Risk Factors for Its Occurrence*, 263 JAMA 1385 (1990). This induced bias is particularly worrisome in meta-analyses which aggregate prior studies in order to draw statistically significant conclusions. Colin B. Begg & Jesse A. Berlin, *Publication Bias and Dissemination of Clinical Research*, 81 J. NAT'L CANCER INST. 107 (1989).

81. Professor Keeton's advice to new judges is simple: "Write opinions rarely." ROBERT E. KEETON, JUDGING 139 (1990).

82. RUGGERO J. ALDISERT, OPINION WRITING 13, 13 (1990).

83. *Id.* at 13-26.

existing law; (iv) constitutes a significant and non-duplicative contribution to legal literature by a historical review of law, by describing legislative history, or by resolving or creating a conflict in the law; (v) reverses a judgment or denies enforcement of an order where the lower court has published an opinion; or (vi) is pursuant to an order of remand from the Supreme Court which is not merely ministerial.[84] Federal district judges also have discretion to request publication of their opinions in the National Reporter System, which selects cases using similar criteria, with the notable—and similarly bias-inducing—addition of cases with unique or unusual fact patterns.[85]

E. Unstated Rationales

Coverage cases are notorious for results-oriented reasoning. A famous quote from Professor Keeton states that "[j]udicial opinions [in coverage litigation] are less than ordinarily enlightening about principled bases for decision. Often . . . the favorite generalization advanced by outside observers to explain a judgment against an insurance company at variance with policy provisions is the . . . aphorism: 'It's an insurance case.'"[86] Spotting ambiguities in policy language (or creating them) is a favorite pastime of judges in coverage cases, as is questioning the impartiality of plan administrators or allowing hindsight to color judgment. For example, an interesting finding of the Hall study was that patients for whom the treatment in question turned out to be effective were twice as likely to prevail in suits to recover damages for the benefit denial as patients treated unsuccessfully. Hall interprets this as indicating that courts are influenced by the unique attributes of cases they consider.[87]

Opinions are written with many audiences in mind.[88] Why a court renders

---

84. *Id.* at 15-17.

85. WEST PUBLISHING CO., PUBLICATION GUIDE FOR JUDGES OF THE UNITED STATES DISTRICT COURTS 2-3 (1994). In addition, on-line services (Lexis and Westlaw) make available various unpublished trial opinions. Unlike unpublished appellate opinions, these may generally be cited as precedent in subsequent litigation.

86. ROBERT E. KEETON, BASIC TEXT ON INSURANCE LAW 341 (1971), *quoted in* Peter Nash Swisher, *Judicial Rationales in Insurance Law: Dusting Off the Formal for the Function*, 52 OHIO ST. L.J. 1037 (1991). Not all commentators are as cynical as Keeton. Swisher, for example, invokes a middle ground between Legal Formalism and Legal Realism to explain judicial reasoning in property and casualty insurance coverage litigation. *Id.* at 1045.

87. Hall et al., *supra* note 2, at 1067. This was true despite the fact that, unlike medical malpractice litigation, causation is not an element of liability in coverage suits. On the other hand, successful treatment implies lower damages, and is often limited to injunctive relief, perhaps making it easier for courts to justify their holdings.

88. According to Leflar, these include posterity, the bar, future judges, the legislature, current and future law students, newspaper readers, the judge himself or herself (to be satisfied with the decision), the parties (especially the loser), and fellow judges (to obtain a majority). Robert A. Leflar, *Some Observations Concerning Judicial Opinions*, 61 COLUM. L. REV. 810, 813-14 (1961). *See also* Ronald A. Cass, *Judging: Norms and Incentives of Retrospective Decision-Making*, 75

a decision and how it explains that decision may therefore differ.[89] Preserving the legitimacy of the judicial system compels reasoning from interpretive principles, while discouraging results-oriented declarations. As Solan notes, there is necessarily a "gap between decision-making and rhetoric in hard cases," although difficult decisions emphasize "seemingly scientific and neutral justification[s]."[90] Supporting this view, surveys of appellate judges frequently yield admissions of conflicts between individual equities or policy considerations and rules of law, although commentators differ as to which holds greater sway.[91] Because some stated rationales are fabrications intended to clothe otherwise naked truth, drawing empirical conclusions from them may be hazardous.

Judges can also avoid hard decisions by retreating into procedural devices. These include mootness, lack of ripeness, lack of adversarialness, non justiciability, lack of standing, failure to exhaust administrative remedies, expiration of limitations periods, or non-compliance with filing requirements.[92] The Hall study deliberately excluded cases which had been resolved on grounds other than the appropriateness of the treatment rendered.[93] However, it is possible that some of the excluded decisions were in fact based on judgments as to appropriateness, but were justified on technical or procedural grounds.

Even if the basis for a decision is not concealed, it may be framed strategically. For example, it is often prudent for a judge to issue as narrow a ruling as possible because the potential consequences of a broader statement are

---

B.U. L. REV. 941 (1995) (examining the structural influences on judges' incentives and behaviors).

89. One piece of evidence for this in coverage litigation is the frequency with which judges disclaim general applicability of their decisions. For example, the court in *Pirozzi v. Blue Cross-Blue Shield* concluded as follows:

> Worth noting here is the modest breadth of this decision. It is not a green light signalling a general expansion of coverage under group health policies like the Plan. Rather, this decision is narrowly, but firmly, anchored in the specific expert medical testimony presented and in the terms and structure of the Plan's experimental exclusion provision. Of course, a different experimental exclusion, or different expert testimony, or a plan that conferred broad discretion on the administrator might well require a different result.

741 F. Supp. 586, 594 (E.D. Va. 1990).

90. LAWRENCE M. SOLAN, THE LANGUAGE OF JUDGES 11, 177 (1993).

91. Llewellyn regarded many judicial constructions as merely providing a means to an already determined end, which he viewed as largely derived from "fireside equities." KARL LLEWELLYN, THE CASE LAW SYSTEM IN AMERICA 79 (Paul Gewirtz ed. & Michael Ansaldi trans., 1989). Marvell, on the other hand, concluded that judges place more emphasis on policy implications. THOMAS B. MARVELL, APPELLATE COURTS AND LAWYERS: INFORMATION GATHERING IN THE ADVERSARY SYSTEM 144 (1978) ("If you can *achieve* justice in that particular case and still do no violence to the law, I'm willing to go along" was a representative comment.). Which set of concerns prevails may differ between trial and appellate courts. *Id.* at 157-58.

92. Posner, *supra* note 41, at 21.

93. Hall et al., *supra* note 2, at 1057.

not knowable at the time.[94] According to Posner, the distinction between holding and dictum buttresses this practice by allowing judges to join opinions with which they do not wholly agree while still not "mortgaging . . . future votes."[95] Similarly, judges frequently prefer to be perceived as constrained in their discretion, and therefore write opinions which portray the court as having but a single option.[96] These proclivities can confound empirical studies which seek to understand the causes underlying judicial outcomes.

IV. WHAT WE MIGHT LEARN FROM JUDICIAL DECISIONS

Despite these limitations, the study of judicial decisions has redeeming qualities. This section describes two ways in which published opinions in coverage cases can help us understand underlying policy issues. One approach is to simplify the empiric inquiry from "what can decisions tell us from their outcomes and reasoning" to "what can decisions tell us from their existence." This avenue can yield information about disequilibrium and adversarialness in the health care system. A second approach is to use judicial opinions to assess the coverage system's ability to bring facts under consideration and to assure fair process. These are essential contributors to the overall success of health insurance, and happen to be things that courts do well.

A. Why Courts Get Involved

Even if we cannot learn as much as we might hope from the content and outcome of reported cases, we can certainly glean information from their incidence. One explanation of the fact that coverage decisions have attracted attention is that the number of coverage disputes generating written opinions has increased markedly over the last decade. The Hall study found that the number of reported cases grew from 5 in the 1960s to 36 in the 1970s, 71 in the 1980s, and 200 in the 1990s.[97] Determining why judicial activity is on the rise may yield important insights into the health care system.

Courts may become active because circumstances are changing and a large number of individuals are aggrieved by the changes. For example, although Llewellyn regards most judicial outcomes as idiosyncratic, he admits that an accumulation of cases favoring one side may induce a shift of the underlying legal rule. He describes this as the result of "a newly emerging consortium of interests pressing hard against an outdated, maladaptive legal norm."[98]

---

94. MARVELL, *supra* note 91, at 223-24.

95. Posner, *supra* note 41, at 20-21. It is not clear how the Hall study treated dicta, or even whether it identified them.

96. SOLAN, *supra* note 90, at 185. Posner calls this the "theory of power without responsibility." Posner, *supra* note 41, at 20.

97. Hall et al., *supra* note 2, at 1060. The last figure was based on a linear extrapolation of cases from 1990 to 1994. Of course, the number of reported opinions in other areas of law has also increased during this period.

98. LLEWELLYN, *supra* note 91, at 100. One might ask whether these interests are

The rapidity of change in the health care system during the recent transition to managed care is self-evident. Major factors include employer-driven cost constraints, federal budgetary retrenchment, and the integration and consolidation of insurance and provider organizations. Earlier transitional periods and their effect on coverage litigation also may be identifiable. For example, increases in reported decisions during the 1970s and 1980s may be related to the impact of new technologies on established underwriting practices and principles of insurance interpretation.

Another possibility is that the health care system is simply becoming more adversarial. In this view, not just the existence of change but its direction promotes litigation. In today's health care system, the erosion of trust produced by the inversion of financial incentives from fee-for-service practice to managed care happens to coincide with a general increase in the aggressiveness of medical consumerism. However, a judicial model of medical decision making is a radical departure from professional traditions in health care, and may have important implications for quality of care and patient satisfaction.

An increase in judicial decision making might also represent an alternative to legislative change. Courts have a recognized role in public policy making.[99] In product liability law, for example, Eisenberg and Henderson speculate that tort reformers failed in their legislative agenda, but still convinced individual judges that reform was needed, as demonstrated by declining plaintiff success rates through the 1980s.[100] In insurance coverage, one wonders whether concerns about cost, or more recently about the excesses of managed care, may have prompted judicial activism during periods of legislative inertia. For example, we know that federal ERISA law has limited state legislative intervention.[101]

In keeping with the earlier discussion of selection and publication bias, however, we should resist the temptation to assume that an increase in reported cases necessarily equates with an increase in underlying disputes. Nonetheless, because many state HMO and insurance regulators now require managed care organizations to maintain records of complaints and grievances, the hypothesis should be verifiable. Moreover, describing the sources of selection, publication, or other biases, should they exist, might be as revealing as confirming their

---

deliberately bringing cases to the attention of courts. Neither malpractice nor coverage has spurred much "impact litigation," in large part because the rewards for individual litigants and their counsel are usually sufficient to ensure frequent judicial review.

99. Mather has identified several aspects of policy making in which courts engage: agenda setting, providing a forum for political argument, agenda building, mobilization of support or opposition, definition of local legal norms, creation of new legal norms, political symbolism and provision of political or legal resources. Mather, *supra* note 69, at 179.

100. Theodore Eisenberg & James A. Henderson, Jr., *Inside the Quiet Revolution in Products Liability*, 39 UCLA L. REV. 731, 751-54 (1992). The authors exclude shifts in accident trends, in the propensity to file claims and in settlement behavior as causes of the decline.

101. ERISA "supersede[s]-any and all State laws insofar as they may now or hereafter relate to any employee benefit plan . . . ." 18 U.S.C. § 1144 (1994).

absence.[102]

B. How Courts Assess Non Judicial Processes

A supportable assertion about reported decisions in coverage cases is that virtually all underwent other levels and forms of review or appeal prior to litigation. Judicial proceedings might therefore shed light on the success or failure of these non judicial processes.

A correlate of the infrequency with which legal precedent is established in coverage litigation is the centrality of facts to the outcome of cases. Although concern about facts is a defining feature of health insurance coverage cases, it is also a staple of litigation in general. In the words of former U.S. Supreme Court Associate Justice Robert H. Jackson,

> It may sound paradoxical, but most contentions of law are won or lost on the facts. The facts often incline a judge to one side or the other. A large part of the time of conference is given to discussion of facts, to determine under what rule of law they fall. Dissents are not usually rooted in disagreement as to a rule of law but as to whether the facts warrant its application.[103]

An important lesson to be drawn from coverage decisions is that fact-finding in modern health care is extremely difficult. For one thing, medical science is generally complex and frequently uncertain. For another, the restructuring of provider organizations in managed care, and the associated financial incentives, have arguably diminished the availability and credibility of information. Therefore, we may be able to learn from judicial decisions how information regarding coverage and care is being shared—or withheld—in managed care organizations. Because accurate, abundant information is central to the long-term success of the health care system, understanding the judicial critique of the mechanism by which information is generated and exchanged in the coverage context could be valuable.[104]

A second lesson relates to procedural fairness. Resource allocation is a critical subtext of insurance coverage litigation. Despite the contractual heritage of health insurance, the litigants and the judicial system are fully cognizant of the social implications of coverage determinations in terms of the cost and equitable distribution of health care.[105] Therapeutic health care is a difficult area for

---

102. These might include, for example, the tension between managed care organizations' concern about the public relations effect of high-profile litigation and the deterrent to settlement produced by limitations on damages under ERISA.

103. MARVELL, *supra* note 91, at 139.

104. There is an important relationship between disclosure in the context of coverage for experimental treatment and medical informed consent. *See* Nancy M.P. King, *Experimental Treatment: Oxymoron or Aspiration*, HASTINGS CENTER REP., July-Aug. 1995, at 6.

105. This sets coverage cases apart from medical malpractice cases. Except for concerns about the cost and efficiency of litigation as a method of dispute resolution, medical malpractice

regulation in large part because lives seem more "identified" than "statistical." In such situations, a finding in favor of coverage allows the question of the marginal value of life to be neatly avoided.[106]

Fairness is a prerequisite to resource allocation, and legal process is the principal guardian of fairness in democratic society. Another important reason to look at judicial decisions as a benchmark for the health care system is therefore that courts are well equipped to evaluate procedural fairness.[107] Judges in coverage cases are suspicious of decisions rendered without due process, and respond favorably to adequate procedural protections for patients and policyholders.[108] Judges' thresholds for procedural fairness are especially high in cases involving preauthorization of services for severe disease, with denial of coverage unlikely unless due process has been scrupulously observed.

We can therefore learn from judicial decisions how well private processes are operating, notably the manner in which coverage standards are developed and the conduct of individual inquiries and appeals. The increase in managed care enrollment by Medicare and Medicaid patients will add to courts' involvement in procedural review,[109] because government programs are subject to more extensive due process requirements than are private parties.[110]

CONCLUSION: CHARTING A RESEARCH AGENDA

A systematic way to approach empirical research is to identify the policy implications of coverage standards and formulate testable hypotheses. In medical malpractice, for example, the objectives of the tort regime are generally

---

is not generally viewed as an issue of resource constraints. Either a treatment was delivered in accordance with the professional standard of care or it was not, and cases are independent of one another.

106. Coverage litigation highlights the distinction between statistical lives and identified lives in our approach to valuing risks. A significant subset of reported coverage decisions involve high-cost therapies for life-threatening conditions. When risk estimated ex ante is converted into loss incurred ex post, and the loss involves human life, it is easy to second guess the earlier valuation.

107. Keeton observes that "good" judges are good because they are skilled at making hard decisions, not because their reasoning is always a model of logic. KEETON, *supra* note 86, at 2.

108. The Hall study did not identify a statistically significant correlation between internal process and case outcome, possibly because of small sample size. *See* Hall et al., *supra* note 2, at 1065-66. The authors have reported elsewhere, however, that courts' objections to internal technology assessment by insurers tended to focus on insufficient or poorly matched sources of information, lack of expert review, concern about financially motivated bias, and lack of current information about clinical benefit. Mark A. Hall et al., *When Courts Review Clinical Practice Guidelines*, MED. CARE (forthcoming 1998).

109. *See* Grijalva v. Shalala, 946 F. Supp. 747 (D. Ariz. 1996).

110. *See* Shelley v. Kraemer, 334 U.S. 1, 13 (1948) (The Fourteenth Amendment "erects no shield against merely private conduct."). However, when a private party's conduct has sufficiently received the imprimatur of the State, it may be deemed state action for purposes of the Fourteenth Amendment. *See, e.g.*, Flagg Bros. v. Brooks, 436 U.S. 149 (1978).

characterized as victim compensation and injury reduction. Health insurance presents a different set of policy concerns. Useful studies will assess the impact of contractual coverage standards (and legislative interventions regarding coverage) on measures such as administrative expenses, health care premiums, medical innovation, patient satisfaction, and health outcomes. For example, the Institute of Medicine estimates that insurers deny only one or two percent of claims, while a much greater amount of care is unnecessary and diverts resources from other areas.[111] Because of the social as well as individual implications of insurance, these are weighty issues.

A threshold question with analogies to medical malpractice is whether courts are reaching efficient and accurate results.[112] For example, Sykes criticizes laws generally allowing "bad faith" claims against insurers because he believes that courts are seldom able to accurately distinguish opportunistic behavior from genuine and reasonable disputes.[113] In health insurance cases, it will be important to assess the degree of correlation or mismatch between valid claims and coverage cases filed, and between valid claims and relief granted.[114] At least for cases involving "medical necessity," some objective scientific determination should be possible. Further studies will explore the cost of grievances and appeals, whether conducted internally to the health plan, through an independent review organization or via the courts.

The pace and direction of research will respond to political constituencies as well as to the interests of academics, and grantmaking bodies will undoubtedly react to both policy and political imperatives. For example, the medical malpractice research agenda of physicians has generally been more focused than that of consumers or attorneys.[115] Consequently, more study has been devoted to the cost of malpractice litigation (groundless claims, administrative expense and defensive medicine), than to quality of care (the amount of substandard

---

111. *See also* Julie A. Jacob, *Managed Care Denials Less Frequent Than Expected*, AM. MED. N., Dec. 15, 1997, at 5 (describing recent studies).

112. There have been simple studies reviewing contractual exclusions and assessing the consistency of treatment of similarly situated individuals, but none has aggregated a large amount of data or drawn statistical rather than anecdotal conclusions. The Hall study included as an explicit goal measuring the potential for inconsistency in judicial decisions, but was hampered by small sample size. Hall et al., *supra* note 2, at 1056, 1058.

113. Alan O. Sykes, *"Bad Faith" Breach of Contract by First-Party Insurers*, 25 J. LEGAL STUD. 405 (1996).

114. *Cf.* Troyen A. Brennan et al., *Relation Between Negligent Adverse Events and the Outcomes of Medical-Malpractice Litigation*, 335 N. ENG. J. MED. 1963 (1996) (finding that severity of disability, not occurrence of adverse events during medical care or related negligence, is predictive of plaintiff recovery for medical malpractice); A. Russell Localio et al., *Relation Between Malpractice Claims and Adverse Events Due to Negligence*, 325 N. ENG. J. MED. 245 (1991) (finding gross mismatch between negligent care and filing of medical malpractice claims).

115. One reason malpractice reform is so important to physicians is because they suffer large psychic damages from litigation which are not compensated by insurance. Kessler & McClellan, *supra* note 52, at 357.

practice and the deterrent effect of litigation) or access to compensation for negligent injury. Political interests in health insurance will probably be dominated by taxpayers concerned about government expenditures under Medicare, large employers seeking to reduce benefit costs and, more likely than not, trial lawyers hoping to exploit the vulnerability of corporate organizations to legal claims. Recent recommendations of the President's Advisory Commission on Consumer Protection and Quality in managed care—including comprehensive procedures for internal and external review of coverage denials—have already prompted advocacy-based research on their likely cost.[116]

Managed care organizations will sponsor technology assessments and cost-effectiveness studies, as will academic health centers and pharmaceutical companies. Employers and other group purchasers will demand statistical proof from insurers and risk-bearing providers that they are receiving value for money, and will probably be compelled under ERISA to communicate this information to beneficiaries. Most importantly, government will mandate reporting by the full range of regulated entities, and will make that information available to researchers. Notably, the expansion of Medicare and Medicaid managed care will expand federal data collection to monitor cost, access and quality in insured systems, and to detect and deter fraud. All of this information will shed light on coverage standards and the processes for making coverage decisions and resolving disputes.

A caveat is that much of this research may not be a planned element of health insurance regulatory design so much as a by-product of data produced for other purposes. This is certainly true in other areas of health law. For example, the existence of comprehensive Medicare data allowed researchers to estimate defensive medicine by linking restrictions on medical malpractice litigation to service utilization.[117]

Although no one can predict exactly how the research agenda will evolve, it is virtually certain that we will witness an extraordinary expansion of empirical work on health insurance coverage over the next few years, much of it based on information extending well beyond judicial opinions. The simple reason is that the stakes—for identifiable constituencies and for society as a whole—are higher than ever before. Higher stakes provoke interest in promoting or resisting change, and greater interest generates data.[118] Our challenge is to interpret those data correctly, and to apply the results responsibly.

---

116. *See Gauging Quality Regulation's Impact on Premium Costs*, MED. & HEALTH, Nov. 24, 1997, at 1; *see also Consumer Bill of Rights and Responsibilities* (visited Nov. 1997) <http://hcqualitycommission.gov>.

117. Kessler & McClellan, *supra* note 52.

118. Of course, there may be political issues, such as abortion, in which the stakes are too high and positions too polarized to admit research.

GETTING A HANDLE ON COVERAGE DECISIONS: IF NOT CASE LAW, THEN WHAT?

Comments on a Paper by Professor William Sage

MAXWELL J. MEHLMAN*

Professor Sage has written an important paper on an important issue. The lack of good empirical work on health law in general, and on the activities of health insurers in particular, is a serious impediment to sound policy making. A good example is the controversy over genetic discrimination in access to health insurance. Much has been written lately about the problem,[1] but no one has any idea of how widespread or serious the problem may be. Without this information, it is difficult to determine the appropriate public policy response.

I am persuaded by Professor Sage's arguments that case law is a poor source of empirical data on coverage decision making by health care insurers. However, I am less optimistic than he appears to be that any good sources of quantitative data on the subject will be forthcoming in the near future. Let me explain why.

First, Professor Sage seems to feel that good empirical data are available in one area of health law, namely, medical malpractice. He states that "medical malpractice data are far more organized and accessible"[2] than data on health insurance coverage, declaring for example, that "data on nationwide jury awards are available from Jury Verdict Research . . . ."[3] However, the Jury Verdict reporting system, which, as Professor Sage recognizes, is based largely on self-reporting by attorneys,[4] is generally regarded as incomplete and unsuitable for quantitative research purposes.[5] Indeed, some of the most basic questions that we might want to ask about the health care system, such as whether malpractice incidence is higher in managed care organizations than in traditional fee-for-service settings, remain unanswerable given the present state of our data.

This makes it all the more discouraging that for coverage decisions we lack data even as good as that we have on malpractice. As Professor Sage recognizes, knowledge of coverage decision making is critical to enable us to determine

---

* Arthur E. Petersilge Professor of Law and Director, The Law-Medicine Center, Case Western Reserve University.

1. *See, e.g.*, Paul R. Billings et al., *Discrimination as a Consequence of Genetic Testing*, 50 AM. J. HUM. GENETICS 465 (1992); Richard A. Bornstein, *Genetic Discrimination, Insurability and Legislation: A Closing of the Legal Loopholes*, 4 J.L. & POL'Y 551 (1996); Kathy L. Hudson et al., *Genetic Discrimination and Health Insurance: An Urgent Need for Reform*, 270 SCIENCE 391 (1995).

2. William M. Sage, *Judicial Opinions Involving Health Insurance Coverage: Trompe L'Oeil or Window on the World?*, 31 IND. L. REV. 49, 59 (1998).

3. *Id.*

4. *Id.*

5. *See, e.g.*, Marc Galanter, *Real World Torts: An Antidote to Anecdote*, 55 MD. L. REV. 1093, 1113 n.63 (1996) ("Because [Jury Verdict Research] does not provide information about the number of awards in its data base, there is no way to tell how thoroughly it represents the entire universe of awards.").

whether patients are obtaining access to new or experimental medical technologies; to identify the criteria third-party payors are using to make these coverage decisions; and, ultimately, to discern how private health care rationing schemes are operating in our present healthcare system.[6] Clearly we need this information. The question is, where and how can we get it?

Professor Sage suggests at the end of his paper that these data will be forthcoming from academic health centers, pharmaceutical companies, government-mandated reports, and the managed care organizations themselves.[7] While I agree with Professor Sage that some additional data will be available from these sources in the future, I am pessimistic that we will obtain easy access to substantial amounts of useful information. To understand why, let me list the types of actions we need to know about in order to have a good understanding of what is happening in the realm of coverage decision making.

1. Whether or Not a Person is Insured.—First, a third-party payor could refuse to cover health care services for an individual by declining insurance to that individual. Therefore, to have a complete picture of coverage, we would need information on insurability decisions by insurers. We also would need to know whether employers offer health care plans to their employees, and, if so, what types of plans they offer. Finally, we would need to know what choices enrollees or potential enrollees make when they are offered plan options by employers or within the insurance market in which they reside. In the Cleveland metropolitan area, for example, employees not offered access to Rainbow Babies and Children's Hospital through their health plan would most likely not have coverage for the specialized services that hospital provides.

2. Premium Levels.—A second way an insurer could decide not to cover services for an individual is to price coverage beyond the individual's ability to pay. Therefore, a complete picture of health care coverage would include knowledge of the premiums associated with different types of plans or coverage options, along with the co-payment and deductible rules for each option.

3. Pre-existing Conditions.—An insurer may try to avoid covering a service by declaring the condition for which the service is required a "pre-existing condition."[8] We would need to know insurers' policies on pre-existing conditions and understand the impact of the Health Insurance Portability and Accountability Act of 1996[9] on those policies.

4. Recommendations by Primary Care Physicians.—One of the most difficult types of coverage decisions to discern is the information primary care physicians give to their patients about the suitability of particular services for the

---

6. Sage, *supra* note 2, at 50.

7. *Id.* at 73.

8. A "pre-existing condition," in general, is one that would "otherwise [be] within the coverage of the policy, [but] which existed prior to [the policy's] effective date." GEORGE J. COUCH, COUCH CYCLOPEDIA OF INSURANCE § 41A:17 (2d rev. ed. 1982). Specific definitions of what constitute "pre-existing conditions" may vary among insurance companies.

9. Pub. L. No. 104-191, 110 Stat. 1936 (1996) (codified in scattered sections of 18 U.S.C., 26 U.S.C., 29 U.S.C., and 42 U.S.C.).

patient, the availability of those services in the geographic area, and coverage of those services under the patient's specific health care plan. Particularly in managed care plans offering financial incentives to physicians for discouraging patients from obtaining expensive health care services, the physicians may make implicit coverage decisions simply by not informing patients about the benefits of a specific service. For example, Henry J. Aaron and William B. Schwartz have shown that physicians in Britain avoided providing kidney dialysis to patients over the age of fifty-five simply by not informing the patients of the availability of dialysis services.[10]

5. Referral Decisions by Gatekeepers.—Managed care organizations increasingly use a system whereby primary care physicians, placed in the role of gatekeepers, are expected to restrict patient access to costly services. For example, the physician may decline to refer the patient to a specialist within the plan, or to a specialist outside the plan, or to a facility within or outside the plan. All of these decisions are denials of access tantamount to negative coverage decisions.

6. Outright Coverage Policy.—Third-party payors typically include coverage language both in their policies and in that portion of the policies provided to enrollees as their description of covered services. To my knowledge, no one has made a comprehensive compilation of policy language from third-party payors around the country. Moreover, much of this policy language is vague, as Professor Sage recognizes in his description of the role of judicial oversight.[11] Thus, one would need to know how the general policy language is interpreted in practice by the health plan. In trying to determine the coverage policy under various state Medicaid programs, for example, we have had to survey the programs directly to determine whether or not they cover a specific technology for a specific patient population. Even so, some responses indicate that to a large extent Medicaid makes coverage decisions on a case-by-case basis regardless of the general statutory language governing Medicaid coverage policy.

7. Utilization Review Decisions.—A primary example of case-by-case interpretations of policy language is utilization review decisions by health plan personnel. These include pre-certification decisions authorizing hospital admissions, extensions of lengths-of-stay, and authorization for specific procedures.

8. Decisions by Plan Medical Directors.—Decisions interpreting general policy language and utilization review decisions by administrative personnel may be reviewed by medical directors when challenged by enrollees or their providers. We need to know the fate of these challenges.

9. Claims Determinations.—Another way in which coverage decision making takes place is through retrospective utilization review. This includes decisions by third-party payors not to reimburse providers and/or enrollees for services already supplied. These refusals color future recommendations by the

---

10. HENRY J. AARON & WILLIAM B. SCHWARTZ, THE PAINFUL PRESCRIPTION: RATIONING HOSPITAL CARE 101 (1984).

11. Sage, *supra* note 2, at 58.

providers to patients and influence future efforts by the same enrollees to obtain similar services. Over time, reimbursement decisions may shape general plan coverage policies.

10. Judicial Review.—Finally, we come to the focus of Professor Sage's paper: cases in which courts review coverage decisions by third-party payors.

Given our need for much or all of the foregoing data to give us an accurate picture of coverage decision making and its effect on enrollees in our health care system, how available is this information at this time? We have some information on who is insured and who is not and some information on employer insurance practices. Although enrollees should have policy language for the plans in which they are enrolled, no one as yet has attempted to gather this policy language in one place, and any effort to do so would be hampered by changing circumstances, such as plans coming into and going out of business, and changes in policy language. We do have judicial opinions, but Professor Sage has done a superb job of persuading us how limited that database is for our purposes.[12]

Given that so much of this information is not available at this time, is Professor Sage correct in predicting that it will become available in the near future?[13] Professor Sage seems to think that we will get a good deal of it from the managed care organizations themselves. But why should they gather this information? They are primarily interested in data that would affect their profitability, and therefore in information about the behavior of providers within their network or providers being considered for membership. Therefore, we might expect them to gather information on provider practice patterns and referral decisions, mentioned in items four and five above. But it is unlikely that health plans would have an interest in gathering much of the other data we desire. Moreover, even if they collected the data on practice patterns and referrals, or indeed any of the other data on our list, why would they make this information public? Assuming they did make it public, how could we verify the accuracy and completeness of the information they disclosed? Even if an isolated managed care organization gave us access to this information, how could we be sure that what we learned from it reflected the practice in other managed care organizations?

The primary force driving the behavior of health care organizations is increasing competition. Should we expect competition to lead third-party payors to publicize their coverage decisions in order to gain a competitive advantage? This is hard to imagine. What health care plan, for example, would want to advertise that it covers an expensive service that other plans do not? Would this not lead to adverse selection by enrollees?

Professor Sage does not rely exclusively on the private sector to provide access to coverage decisions. In addition, he states that "government will mandate reporting by the full range of regulated entities, and will make that information available to researchers."[14] Again, I fear this is too optimistic. The

---

12. *Id.* at 61-68.
13. *Id.* at 73.
14. *Id.*

one potential source of good malpractice data is the National Practitioner Data Bank and it is true that this database has been created by the federal government. But the data are not accessible to researchers, let alone to the public.[15] Nor does it appear that the government is making a systematic effort to ensure that reporting to the National Practitioners Data Bank is accurate and complete. Even if the government were to require similar reporting for coverage decisions, we would be hard pressed to ensure that they too were accurate and complete. Furthermore, would the government take this step when doing so would increase the administrative burden on third-party payors? If the answer is that medical records will be computerized and therefore much easier to retrieve and analyze, does this not raise privacy questions that might block government data collection efforts? Finally, how could the government obtain information on referral practices and recommendations by primary care physicians? The only way would seem to observe physician office practice directly, a solution either impractical, unethical, or both.

So far I have only been talking about getting information from third-party payors themselves. Another potential source of information might be enrollees. Some studies of discrimination and access to health care services, for example, have relied on self-reporting by enrollees or potential enrollees.[16] But it is well recognized that patient self-reporting is inaccurate, biased and incomplete. An interesting possibility is to identify and contact enrollees to get their permission to access their medical records, and then to analyze the records to ascertain what coverage decisions may have been made in their cases. Gerald Hickson and his associates have used a similar technique to determine why patients who have arguably been the victims of medical malpractice file claims.[17] But this promises to be an extremely expensive and laborious process, and it is unclear whether sufficient information could be ascertained by this method.

In conclusion, I want to stress two points. First, the difficulty of obtaining data in no way diminishes our need for it. Second, I think we are going to be limited to more theoretical, extrapolative efforts to examine coverage decisions for the foreseeable future.

---

15. 42 U.S.C. § 11137 (1994).

16. *See* E. Virginia Lapham et al., *Genetic Discrimination: Perspectives of Consumers*, 274 SCIENCE 621 (1996).

17. Gerald B. Hickson et al., *Factors that Prompted Families to File Medical Malpractice Claims Following Perinatal Injuries*, 267 JAMA 1359 (1992).

EMPIRICAL STUDIES OF JUDICIAL DECISIONS SERVE AN IMPORTANT ROLE IN THE CUMULATIVE PROCESS OF POLICY MAKING

Comments on a Paper by Professor William Sage

KAREN A. JORDAN*

INTRODUCTION

Professor Sage has presented a cogent analysis of the role that empirical studies of judicial opinions can play in the shaping of health law and policy.[1] For those of us who have an interest in empirical research, but who are not as well versed as we would like to be in statistics or the basics of social science research, the most beneficial aspect of Professor Sage's article is the section detailing the limitations associated with using judicial decisions as an empirical data set.[2] My comments are largely directed to that aspect of his article. Overall, I agree with Professor Sage's conclusion that, despite the limitations of using decisions as a data set, empirical studies of judicial decisions can yield useful information. However, this commentary stresses that the limitations of studying judicial decisions are not necessarily as critical as suggested, and that empirical findings of such studies can play an important role in the cumulative process of policy making.

I. SOME FUNDAMENTALS OF EMPIRICAL RESEARCH

To assess the seriousness of the limitations associated with the empirical study of judicial decisions, it is useful to first review some fundamentals of empirical research designed to facilitate social policy.[3] The function of policy research is to generate information that can render policy making more effective. Broadly speaking, then, the goal of empirical research in this context is to verify propositions about some aspect of the relationship between the objectives of the policy and the means available to achieve those ends.[4] However, effective policy making is a process. For example, policy making has been explained as a series

---

* Associate Professor of Law, Louis D. Brandeis School of Law at the University of Louisville. The author formerly served as Assistant Director of the Center for Law and Health and congratulates the Center on its Tenth Anniversary.

1. William M. Sage, *Judicial Opinions Involving Health Insurance Coverage: Trompe L'Oeil or Window on the World?*, 31 IND. L. REV. 49 (1998).

2. *Id.* at 57-64.

3. Several texts provide a detailed review of empirical research design and implementation. *See generally* OTTAR HELLEVIK, INTRODUCTION TO CAUSAL ANALYSIS: EXPLORING SURVEY DATA BY CROSSTABULATION (1984); ROBERT J. MUTCHNICK & BRUCE L. BERG, RESEARCH METHODS FOR THE SOCIAL SCIENCES: PRACTICE & APPLICATIONS (1996); CAROL H. WEISS, USING SOCIAL RESEARCH IN PUBLIC POLICY MAKING (1977).

4. ROBERT R. MAYER & ERNEST GREENWOOD, THE DESIGN OF SOCIAL POLICY RESEARCH 57 (1980) (describing interaction as a "means-end" relationship).

of stages which includes the following: the determination of goals, needs assessment, specification of objectives, design of alternative courses of action, estimation of consequences of alternative action, selection of courses of action, implementation, and evaluation of outcomes.[5]

To put this into context, consider the problem faced by persons who receive their health coverage through plans governed by the Employee Retirement Income Security Act of 1974 (ERISA).[6] Because of the interplay between ERISA's civil enforcement provisions and ERISA's preemption provisions,[7] policyholders who sustain injury as a result of a denial of coverage by the plan administrator are often unsuccessful in challenging those denials in court.[8] Those interested in health law and policy may therefore determine that some modification of the law of ERISA is warranted to ensure that claimants covered through ERISA plans receive equitable treatment when forced to bring their disputes into the judicial system.[9] If so, the next stage in the policy-making process is to perform a needs assessment. That is, a policy maker should seek information confirming the need for a change in policy and about the amount of change in existing law required to reach the goal.[10] Descriptive research provides information relevant to this stage of the process.

Descriptive research studies are designed to yield quantitative measurements of the characteristics (also referred to as properties or variables) associated with the phenomenon under study.[11] For example, to assess the need for change and the amount of change required in the ERISA context, a study might simply measure how "the law applied" affects the outcome of the case; in other words, did the claimant win less often if the court applied ERISA than when state contract law was applied.[12] Descriptive studies yield information such as the

---

5. *Id.* at 58.

6. 29 U.S.C. §§ 1001-1461 (1994 & Supp. I 1995).

7. ERISA's civil enforcement provisions permit certain suits to be brought against benefit plans by plan participants and beneficiaries but limit the remedies available. *See* 29 U.S.C. § 1132(a) (1994). ERISA's preemption provisions provide that ERISA supersedes state laws that "relate to" ERISA plans, unless exempted as a law that regulates insurance, banking or securities. *See* 29 U.S.C. § 1144(a), (b)(2)(A) (1994 & Supp. I 1995).

8. *See, e.g.*, Tolton v. American Biodyne, Inc., 48 F.3d 937 (6th Cir. 1995) (holding preempted a claim arising out of continued denial of request for inpatient mental care which led to the patient's suicide); Corcoran v. United Healthcare, Inc., 965 F.2d 1321 (5th Cir.), *cert. denied*, 113 S. Ct. 812 (1992) (holding preempted a claim for wrongful death arising out of a denial of coverage for inpatient care for high risk pregnancy).

9. *See, e.g.*, Karen A. Jordan, *Travelers Insurance: New Support for the Argument to Restrain ERISA Pre-emption*, 13 YALE J. ON REG. 255 (1996).

10. MAYER & GREENWOOD, *supra* note 4, at 10.

11. *Id.* at 54.

12. *See* Mark A. Hall et al., *Judicial Protection of Managed Care Consumers: An Empirical Study of Insurance Coverage Disputes*, 26 SETON HALL L. REV. 1055, 1062 (1996). The study was analyzed by Professor Sage and included this type of inquiry.

size, distribution, and interrelationships between variables.[13]

Next, the policy maker should identify the means available for achieving the goal of fair outcomes. This inherently involves an assessment of which factors lead to unjust results in ERISA cases. Exploratory research methods are appropriate for this stage of the policy-making process. Exploratory research involves the careful selection of a few units, which are studied comprehensively by means of a variety of unstructured and unrefined data collection techniques.[14] The data are analyzed qualitatively, rather than quantitatively, for certain inferences about, for example, what factors amenable to judicial intervention seem to be associated with unjust results in ERISA cases. In the ERISA context, exploratory research would likely lead to the inference that factors such as an abuse of discretion standard of review, exhaustion of administrative remedies, or seeking damages beyond the benefit itself are relevant for further study.[15] However, further descriptive research would be necessary to determine the size, distribution and interrelations between these factors or variables. Additionally, research designed to assess the causal relation between each variable and the outcome would help the policy maker ascertain what change in the law of ERISA would most likely lead to more just outcomes.[16]

Empirical research of judicial decisions would appear to be the obvious unit of study for each of the described stages—the descriptive research performed at the needs assessment stage, the exploratory studies performed to identify the relevant factors or variables, and the further descriptive studies to quantitatively assess the size, distribution, interrelations and causal connections between the relevant variables. Moreover, due to the existence of a readily accessible data set, a researcher would likely combine all of these steps into one research study. The question is whether the limitations associated with the use of judicial decisions as a data set would compromise such a research study.

---

13. MAYER & GREENWOOD, *supra* note 4, at 54-55. Studies using the descriptive method are either univariate, describing the distribution of a single variable, or multivariate, describing the simultaneous distribution of two or more variables. *Id.* at 55.

14. *Id.* at 52-53.

15. *See, e.g.*, Barnett v. Kaiser Found. Health Plan, Inc., 32 F.3d 413 (9th Cir. 1994) (affirming denial of coverage for a liver transplant under the abuse of discretion standard applicable to the decision of the ERISA plan administrator); Corcoran v. United Healthcare, Inc., 965 F.2d 1321 (5th Cir. 1992) (dismissing the ERISA claim because plaintiffs sought damages for emotional distress and mental anguish, remedies not prescribed by section 502(a) of ERISA). *See also* Hall et al., *supra* note 12, at 1062.

16. The most sophisticated form of research, explanatory research, is generally not used in policy research until after a change in policy has been implemented. At this point, an evaluation of outcomes helps assess whether the policy change promoted the stated policy goal. The focus of a study undertaken for this purpose is whether the change in policy (the means) has a causal relationship with the objective (the ends). MAYER & GREENWOOD, *supra* note 4, at 59-60.

II. ASSESSMENT OF THE LIMITATIONS

Professor Sage discussed limitations stemming from small sample size, time lags, and selection bias.[17] As he noted, the time lag aspect does not affect the quality of an empirical study of decisions, but only its effectiveness, or relevance, in a rapidly changing market.[18] However, small sample size and selection bias could impact the quality of research findings. The quality of findings in empirical studies is generally assessed with respect to four criteria: their generalizability, their validity, their reliability, and their practical significance.[19] The limitations pointed out by Professor Sage relate primarily to generalizability and validity. Accordingly, these concepts are explored before assessing whether the limitations of sample size and selection bias compromise empirical studies of judicial decisions.

"Generalizability" refers to the extent to which empirical findings can be generalized as representative of cases other than those studied.[20] The term "external validity" is also sometimes used to describe this aspect of a study.[21] Generalizability is relevant when the unit selected to be studied (for example, judicial decisions) consists of a large number of those units, and, because it is not possible to study all of them, a subset of the group is selected that hopefully represents the larger body. However, when the units selected for study do not correspond exactly to the target population (defined to mean the aggregation of units to which the study findings are hoped to apply),[22] generalizability is compromised.[23]

Although some of the selection bias problems identified by Professor Sage as inherent in a data set comprised of judicial decisions would affect generalizability, some would not. Professor Sage pointed out that, among other things, reported decisions represent the tip of the iceberg—they are not the average case and thus are not particularly representative of how most patients fare in their out-of-courtroom coverage disputes.[24] However, this form of selection bias would not necessarily be detrimental to an empirical study of

---

17. Sage, *supra* note 1, at 61-65.

18. *Id.* at 62.

19. MAYER & GREENWOOD, *supra* note 4, at 257.

20. *Id.*

21. *See, e.g.*, JOHN W. CRESWELL, RESEARCH DESIGN: QUALITATIVE & QUANTITATIVE APPROACHES 158-59 (1994); CURTIS D. HARDYCK & LEWIS F. PETRINOVICH, UNDERSTANDING RESEARCH IN THE SOCIAL SCIENCES: A PRACTICAL GUIDE TO UNDERSTANDING SOCIAL AND BEHAVIORAL RESEARCH 7 (1975).

22. MAYER & GREENWOOD, *supra* note 4, at 170-72. For example, consider a hypothetical ERISA study designed to determine whether, how, and to what extent the law of ERISA should be modified to ensure equitable treatment in the judicial system. The target population would include all persons covered through ERISA plans who look to the courts for some consumer protections. The researcher has access to the target population through judicial decisions.

23. The study findings can be applied only to those members of the target population which also fall into the study population. *Id.* at 258.

24. Sage, *supra* note 1, at 58.

judicial decisions. For example, the Hall Study was designed to assess the level of consumer protection available through judicial review of coverage disputes.[25] That is, the study focused on how courts enforce contractual entitlements to health benefits. Thus, in the Hall Study, the fact that the cases studied would not be representative of how patients fare in their out-of-court disputes would not have a negative impact on study results. The empirical findings would be generalizable to the whole target population—those who resort to the judicial system.

A second form of selection bias noted by Professor Sage similarly has a minimal effect on the generalizability of the Hall Study findings. Extrapolating from the low number of cases involving self-insured plans, Professor Sage opined that ERISA cases may have been underrepresented and further suggested that the likelihood of preemption of the claim by ERISA is a powerful deterrent to suit.[26] Again, however, because the focus of the Hall Study was on judicial treatment of coverage disputes, this effect on the number of judicial decisions would not affect generalizability. Rather, the effect simply highlights the narrowness of the study. The Hall Study findings are indicative of judicial treatment of coverage disputes; they are not indicative of the way ERISA plan participants fare in the administrative treatment of coverage disputes. Although this aspect of selection bias highlights the narrowness of the data produced when judicial decisions are studied in empirical research, it has a minimal effect on generalizability.

However, two other selection bias problems would appear to have a more substantial impact on generalizability. Professor Sage explained that publication bias exists because judges have discretion in deciding which of the opinions they write will be reported.[27] Similarly, Professor Sage pointed out that judges have discretion in deciding which of the many factors that may have influenced their decision will appear in the written opinion and that, "[b]ecause some stated rationales are fabrications intended to clothe otherwise naked truth, drawing empirical conclusions from them may be hazardous."[28] These limitations could impact generalizability because they suggest that the findings would not accurately reflect how other patients would fare in cases taken to court.

Perhaps more importantly, the limitations stemming from judicial discretion would also have an impact on the validity of a study's empirical findings. Validity (specifically, internal as opposed to external validity) concerns the accuracy of empirical findings and whether they match reality.[29] Thus, validity refers to the extent to which the study findings are applicable or relevant to the research objectives and, more specifically, to the extent to which the measures obtained reflect the variables specified in the research objectives.[30]

---

25. Hall et al., *supra* note 12, at 1056.

26. Sage, *supra* note 1, at 65.

27. *Id.* at 65-66.

28. *Id.* at 67.

29. CRESWELL, *supra* note 21, at 158.

30. MAYER & GREENWOOD, *supra* note 4, at 258. The principle limitation to validity is often the indicators selected to measure concepts intended to be studied. For example, if a study

The problem of unstated rationales in judicial opinions would arguably impact the validity of a study of judicial opinions. For example, in our hypothetical ERISA study, one variable studied might be the use of the abuse of discretion standard of review. If a judge's opinion was written so that it appeared that the standard of review was key to the outcome but there were other unstated rationales, the measure of that variable would not match reality. Similarly, if the data set is underrepresentative of ERISA cases, as Professor Sage suspected in the Hall Study, this would impact the validity of the finding that ERISA does not have a statistically significant association with a claimant's outcome.[31]

The other major limitation identified by Professor Sage was sample size. He stated that there are far too few reported decisions pertaining to coverage disputes to draw statistically meaningful conclusions.[32] In the Hall Study, however, sample size became a problem because the study was limited to coverage disputes involving denials due to medical appropriateness, defined to include decisions turning on medical necessity, or on whether the treatment could be characterized as experimental or investigational.[33] This focus reduced the sample size by over 750 cases.[34] Thus, it is difficult to predict the extent to which small sample size may or may not be a problem in studies of other types of coverage disputes. For example, we are seeing more cases involving challenges to precertification procedures that delay care, or to denials stemming from unreasonable financial incentives to reduce care.[35] Studies of other coverage issues might well result in a sufficiently sized sample.

But sample size is important. Professor Sage explained that small sample size can limit the choice of data analysis techniques.[36] For example, descriptive research studies designed to yield information such as the interrelationships between variables usually require a large number of units of study.[37] Further, size can impact validity, reliability and generalizability of empirical findings.[38]

---

used age and level of education as indicators of a person's employability, validity would depend on degree of association between age and level of education and the length of time required to secure a job. *Id.* at 258-59.

31. *See* Hall et al., *supra* note 12, at 1066.

32. Sage, *supra* note 1, at 61.

33. *See* Hall et al., *supra* note 12, at 1057.

34. *Id.* at 1057-58.

35. *See, e.g.*, Pappas v. Asbel, 675 A.2d 711 (Pa. Super. 1996), *appeal granted*, 686 A.2d 1312 (1996) (involving allegedly negligent delay in pre-authorization); Ouellette v. Christ Hosp., 942 F. Supp. 1160 (S.D. Ohio 1996) (involving an allegedly unreasonable system of financial incentives that caused premature discharge from the hospital).

36. Sage, *supra* note 1, at 61.

37. MAYER & GREENWOOD, *supra* note 4, at 54.

38. Reliability refers to the degree of confidence that can be accorded the research findings, generally considered to be the extent to which repeated applications of the research design under similar conditions would yield consistent findings. *Id.* at 259. Reliability may be hindered in two principle ways: if the study population is selected randomly, or if there is variability in the data collection techniques. However, in both instances, the degree of reliability may be measured. *Id.*

This is because of the fundamental role that probability plays in statistical processes. Because researchers hope to prove that their findings represent more than a chance relationship between variables, probability is central to statistical significance and sampling.[39] For example, if the study population is selected randomly, there must be a sufficient number of units selected such that there is a high probability of reproducing the essential characteristics of the total population.[40] However, there is not a generally recognized "requisite size" for an empirical study. Rather, a researcher generally must balance the need for reliability and validity against the costs and benefits of the study.[41]

In my view, the need for balancing costs and benefits is the key to the question of whether judicial decisions should be used as a data set for empirical research. There are real limitations associated with an empirical study of judicial decisions. However, those limitations must be balanced against the benefits that can be gained from the information revealed. Indeed, some empirical researchers have concluded that validity is a concept "to be pursued, but not to be attained."[42] They have rejected the common view that if researchers "can acquire a sufficient amount of [validity], by applying appropriate techniques, one has somehow 'won' at the game called research."[43] Rather, they advocate that validity is to be assessed "relative to purposes and circumstances."[44]

Thus, although there are limitations associated with studying judicial decisions, those limitations must be considered in light of the reasons researchers might want to study judicial decisions. This comment has already explained that empirical studies of judicial decisions may reveal information useful in various stages of the policy making process. For example, such studies have been found helpful in assessing the need for a change in policy or in formulating possible alternatives.[45] Further, Professor Sage pointed out that studies of judicial decisions can provide useful information. For example, studies can provide information about the increasing number of cases ending up in court and possibly some insight into the reasons for the increase or information about the administrative processes which patients must exhaust prior to seeking redress in

---

39. KENNETH R. HOOVER, THE ELEMENTS OF SOCIAL SCIENTIFIC THINKING 97-98 (4th ed. 1988).

40. *Id.* at 99.

41. MAYER & GREENWOOD, *supra* note 4, at 178 (noting that the policy researcher tries to increase the number of observations in order to maximize reliability, but that the importance of reliability will vary from one type of study to another).

42. DAVID BRINBERG & JOSEPH E. MCGRATH, VALIDITY AND THE RESEARCH PROCESS 13 (1985).

43. *Id.*

44. *Id.* (emphasis omitted). The authors further note that rather than viewing validity as a necessary aspect of an empirical study, it would be more productive to pursue robustness analyses of empirical findings, i.e., further activities designed to assess the degree of certainty surrounding the findings. *Id.* at 119-38. They specify three sets of activities to assess robustness: replication, convergence analysis, and boundary search. *Id.* at 136.

45. *See supra* notes 1-17 and accompanying text.

court.[46]

These examples highlight the fact that empirical research is, by nature, progressive. Each study yields discrete pieces of information relevant to overarching policy concerns.[47] Thus, studies of judicial decisions yield useful, albeit narrow information, that moves us toward a greater understanding of the bigger policy questions. For example, the Hall study was designed to assess consumer protection provided by courts. It was therefore designed to determine, from an empirical perspective, what factors are associated with judicial enforcement of a contractual right to medically necessary benefits. Specific factors were selected for study, including, for example, the type of insurance (public programs or private or government employer); what law governs (state contract law, state statute, federal statute, ERISA); and discretion assigned to the insurer.[48] The study was designed therefore to ascertain small, but key points. For example, do those who obtain coverage through public programs prevail significantly less often than those covered by private insurance or government employees? The study showed that in fact they prevailed in seventy percent of the cases; and further, that government employees prevailed in only thirty-one percent of the cases.[49] This empirical finding could lead to research to determine why government employees prevail less often.

The Hall study also showed that when the dispute is governed by ERISA, the patient/policyholder is less likely to win.[50] Individuals familiar with ERISA cases would have guessed that this was the case. Nonetheless, this empirical finding is important because it substantiates and quantifies the information. Thus, the Hall study, as well as other studies of judicial decisions, are likely to produce important pieces of empirical information despite the fact that the data set is far from perfect. The information may be important in and of itself. However, it is also important because it provides direction for further research. Thus, to answer the question Professor Sage presented, empirical analysis of judicial decisions does *not* risk the absurdity of "looking for the lost coin under the lamppost solely because the light is better."[51]

CONCLUSION

Two further comments are noteworthy. First, although I have concluded that empirical studies of judicial decisions are a worthy pursuit despite their limitations, I recognize that it is the additional empirical research that stems from

---

46. Sage, *supra* note 1, at 68.

47. HARRIS M. COOPER, INTEGRATING RESEARCH: A GUIDE FOR LITERATURE REVIEWS 11 (2d ed. 1989) (explaining that, because of the cumulative nature of science, trustworthy accounts of past research form a necessary condition for orderly knowledge building).

48. Other variables included coverage language (general or specific); jurisdiction (state appeals, federal trial, federal appeals); likelihood of death; and seriousness of patient's condition. *See* Hall et al., *supra* note 12, at 1066-67.

49. *Id.* at 1062.

50. *Id.*

51. Sage, *supra* note 1, at 50.

studies of judicial decisions that will most likely shape health law and policy. For example, in the ERISA context, the Hall study revealed that patients covered through ERISA plans, and thus whose disputes arising from a denial of coverage must be pursued as an ERISA claim, are *less* likely to win in court.[52] Although this is important information, there is still a larger empirical issue. Namely, does that outcome promote or hinder the policy goals underlying ERISA. Further empirical studies are needed to assess this broader policy question.

For example, the primary policy goal underlying ERISA, and specifically underlying ERISA's civil enforcement provisions which allow suits, but limit the available remedies,[53] is that ERISA plans should be shielded from the *financial risk* associated with a denial of coverage in order to protect the plan as a whole. Thus, ERISA plans are protected from punitive damages and from compensation beyond the "benefit" itself, in order to protect the plan as a whole.[54] The important empirical question, then, is whether the goal of protecting the plan is promoted. Accordingly, studies should be designed to assess, among other things, the impact of fewer patients winning on variables such as administrative expenses, premiums, health outcomes or patient satisfaction. A study designed to test whether the fact that patients prevail less often results in lower administrative expenses or health care premiums would need to compare administrative expenses and premiums in ERISA plans with expenses and premiums in non-ERISA plans. Such information is far beyond the information found in judicial decisions.

More sophisticated empirical studies such as these will more readily facilitate the shaping of health law and policy. For example, as a legal scholar I have devoted substantial time and energy to developing doctrinal arguments that can be used to take cases outside the scope of ERISA,[55] because I believe the law of ERISA should be modified so that patients covered through ERISA plans are treated comparably to those in non-ERISA plans. However, an empirical analysis showing that, although fewer patients win their benefits claims under current ERISA regulations, the goal of greater benefit to the plan as a whole is advanced might convince me that the law of ERISA does not need to modified. Thus, although empirical studies of judicial decisions bring important information to light and are important initial steps that point out the direction of further research, it is the further research which is more likely to be influential in shaping health law and policy.

However, the influence of empirical studies may be limited in the judicial

---

52. The Hall research team speculated that this was due to the standard of judicial review in ERISA cases when the insurer has been granted discretion. *See* Hall et al., *supra* note 12, at 1062-63.

53. *See supra* note 7.

54. *See, e.g.*, Massachusetts Mutual Life Ins. Co. v. Russell, 473 U.S. 134 (1985) (holding that extra-contractual and punitive damages were not available in a claim under 29 U.S.C. § 502(a)(1)(B) (1994) of ERISA and the plaintiff was limited to the remedies set forth in § 502(a)).

55. *See, e.g.*, Jordan, *supra* note 9, at 255; Karen A. Jordan, *ERISA Pre-emption: Integrating Fabe into the Savings Clause Analysis*, 27 RUTGERS L.J. 273 (1996).

arena. This is because empirical research is, in some ways, at odds with legal analysis. That is, although empirical studies might influence health law scholars, they are less influential to lawyers in the position of advocate. Again, consider coverage disputes in the context of ERISA. An attorney in private practice faced with a client with a coverage dispute may discount empirical findings. For example, assume the following facts: 1) a client needed cardiac surgery immediately; 2) her HMO first required a second opinion; 3) her HMO then disagreed about what facility should be used; and 4) the client subsequently did not obtain the requisite preauthorization until her condition had deteriorated to the extent that surgery was no longer a viable option. In this case, an attorney would likely believe that the HMO was negligent and would want to find a way to hold the HMO responsible. This tactic would be the appropriate way to attain compensation for the client and her family and to prompt the HMO to develop expedited precertification procedures so that this unfortunate situation would not occur again.

However, based on current case law, it would be difficult to hold the HMO liable if the client obtained her coverage through a plan governed by ERISA.[56] Nonetheless, to serve the client, the attorney would engage in a "legal analysis" of the judicial decisions, looking for ambiguities or inconsistencies in the case law, or distinguishing specific facts. The attorney would find a meritorious legal argument that would support imposing liability on the HMO, because the job of advocate requires it, and because the process of legal analysis permits it.[57] Moreover, an advocate would do this even though empirical studies may exist showing that holding the HMO liable would hinder public policy and would be worse for the ERISA plan than not pursuing the case at all.

This type of scenario highlights the fact that legal analysis and empirical analysis may often be like ships that pass in the night. In the ERISA context, and perhaps in others as well, they seem to serve cross purposes. One possible consequence of this is that, although empirical studies *should* help shape health law and policy, it is less likely that they will have a great impact in the judicial arena. This lack of impact is because the judicial system is driven by individual disputes; arguments are advanced by attorneys in their role as advocate; and judges, although attentive to empirical data, will ultimately be striving to ensure justice in the individual case before the court.

---

56. Most courts readily find that cases involving allegedly negligent denials of coverage are preempted by ERISA. *See, e.g.*, Jass v. Prudential Health Care Plan, Inc., 88 F.3d 1482 (7th Cir. 1996); Kuhl v. Lincoln Nat'l Health Plan of Kansas City, 999 F.2d 298 (8th Cir. 1993), *cert. denied*, 510 U.S. 1045 (1994); Corcoran v. United Healthcare, Inc., 965 F.2d 1321 (5th Cir.), *cert. denied*, 506 U.S. 1033 (1992).

57. *See, e.g.*, LINDA HOLDEMAN EDWARDS, LEGAL WRITING: PROCESS, ANALYSIS AND ORGANIZATION (1996); ROBERT E. RODES, JR. & HOWARD POSPESEL, PREMISES AND CONCLUSIONS: SYMBOLIC LOGIC FOR LEGAL ANALYSIS (1997).

THE APPLICATION OF ANTITRUST DOCTRINE TO THE HEALTHCARE INDUSTRY: THE INTERWEAVING OF EMPIRICAL AND NORMATIVE ISSUES

JAMES F. BLUMSTEIN*

INTRODUCTION

At present, much of healthcare policy is dominated by a debate about different ways of thinking about medical care—about different paradigms. On one hand, proponents of the traditional professional paradigm have argued, empirically, that the market cannot work well in medical care and, normatively, should not be permitted to work, at least in some situations.[1] Market-skeptics typically contend that medical care involves technical decisions that are beyond the ability of consumers to make. Professionals, with much training and a claim to scientific expertise, should be entrusted with medical care decision making because of what is characterized as the asymmetry of information—providers have it, consumers do not. In this account, consumers lack the knowledge to participate actively in decisions regarding medical care and, presumably, are incapable of becoming adequately informed, either directly or through use of information intermediaries.[2] Proponents of the professional paradigm claim that medical decision making is basically scientific, made by autonomous professional providers. They also claim that economic incentives should not significantly affect these professional, scientific decisions.[3]

In the last fifteen years, the normative and empirical premises of the professional paradigm have come under challenge. Patients have demanded to

---

* Professor of Law, Vanderbilt Law School. Director, Health Policy Center, Vanderbilt Institute for Public Policy Studies. Work on this article was supported in part by a grant from the Robert Wood Johnson Foundation.

1. *See, e.g.*, Arnold S. Relman, *Practicing Medicine in the New Business Climate*, 316 NEW ENG. J. MED. 1150 (1987) (arguing that "the present trend toward market competition is clearly weakening values of our profession.").

2. Thomas L. Greaney, *Quality of Care and Market Failure Defenses in Antitrust Health Care Litigation*, 21 CONN. L. REV. 605, 633-35 (1989). This results either from incapacity or because of the high costs of becoming informed. For a classic description and analysis of the purported market failure in medical care, see Kenneth J. Arrow, *Uncertainty and the Welfare Economics of Medical Care*, 53 AM. ECON. REV. 941 (1963). *See also* PAUL STARR, THE SOCIAL TRANSFORMATION OF AMERICAN MEDICINE 226-30 (1982) (arguing that professionalism, which Arrow regards as a response to market failure, is to some extent a cause of some market problems).

3. A classic statement of this adherence to the professional paradigm is Arnold S. Relman, *The New Medical-Industrial Complex*, 303 NEW ENG. J. MED. 963 (1980); *see also* Arnold S. Relman, *Medical Practice Under the Clinton Reforms—Avoiding Domination by Business*, 329 NEW ENG. J. MED. 1574 (1993). For a discussion of the professional culture and its hostility to financial incentives, see David M. Frankford, *Managing Medical Clinicians' Work Through the Use of Financial Incentives*, 29 WAKE FOREST L. REV. 71 (1994).

participate in decisions involving their own medical care,[4] and physicians increasingly seem to recognize the clinical significance of expanded patient participation in medical decisions affecting patients' own lives and health.[5] Studies have shown that there is a wide divergence in patterns of utilization among providers.[6] Perhaps the best-known of these studies was published in 1996 as the Dartmouth Atlas,[7] which demonstrated the wide variations of utilization in different regions without any ostensible scientific rationale. Further, earlier Rand experiments showed that financial disincentives significantly affected consumers' behavior,[8] and studies associated with the introduction of Diagnosis Related Groups (DRGs) in Medicare in the mid-1980s showed a considerable impact on hospital length of stay.[9]

The dramatic shift to outpatient medicine has undoubtedly been influenced by the incentives associated with DRGs. Evidence has also shown that different patterns of practice are associated with managed care (as contrasted with traditional fee-for-service medicine).[10] In short, it became clear that medical care, like other economic goods and services, responded to the economic realities and incentives of the marketplace.[11] The pure professional paradigm, which initially rejected such economic effects as a matter of empirical totem, has come to resist as a normative matter what is now seen as corruptive of medical practice.[12] Ignoring such economic realities, however, has resulted in extraordinary cost escalation when linked with third-party insurance, and has had

---

4. *See generally* Peter H. Schuck, *Rethinking Informed Consent*, 103 YALE L.J. 899 (1994).

5. *See* Joseph F. Kasper et al., *Developing Shared Decision-Making Programs to Improve the Quality of Health Care*, QUALITY REV. BULL., June 1992, at 183, 184.

6. *See, e.g.*, John E. Wennberg, *Dealing with Medical Practice Variations: A Proposal for Action*, HEALTH AFF., Spring 1984, at 6; Mark R. Chassin et al., *Variations in the Use of Medical and Surgical Services by the Medicare Population*, 314 NEW ENG. J. MED. 285, 287-89 (1986).

7. THE CENTER FOR THE EVALUATIVE CLINICAL SCIENCES, DARTMOUTH MEDICAL SCHOOL, THE DARTMOUTH ATLAS OF HEALTH CARE (1996).

8. *See generally* JOSEPH P. NEWHOUSE ET AL., FREE FOR ALL? LESSONS FROM THE RAND HEALTH INSURANCE EXPERIMENT (1993).

9. *See generally* Judith R. Lave, *The Impact of the Medicare Prospective Payment System and Recommendations for Change*, 7 YALE J. ON REG. 499 (1990).

10. *See generally* Robert H. Miller & Harold S. Luft, *Managed Care Plan Performance Since 1980*, 271 JAMA 1512 (1994). For a discussion of physician attitudes towards managed care, see Gail Silverstein, *Physicians' Perceptions of Commercial and Medicaid Managed Care Plans: A Comparison*, 22 J. HEALTH POL., POL'Y & L. 5, 9-10 (1997).

11. *See* James F. Blumstein & Frank A. Sloan, *Redefining Government's Role in Health Care: Is A Dose of Competition What the Doctor Should Order?*, 34 VAND. L. REV. 849, 852 (1981).

12. For an expression of this view in the context of a determination of liability, see *Muse v. Charter Hosp. of Winston-Salem, Inc.*, 452 S.E.2d 589 (N.C. Ct. App.), *aff'd per curiam*, 464 S.E.2d 44 (N.C. 1995).

to be reassessed from a policy perspective.[13]

Serious analysts do not call for the elimination of the professional paradigm. Rather, what is needed and what has occurred is an accommodation between elements of both models.[14] The pure professional paradigm is no longer feasible or appropriate, but the professionalism of physicians and other medical care providers is still an important component of the high quality of American medicine. While economics cannot be ignored, neither should the important professional contributions of medical care providers be underestimated or undervalued.[15]

An important contributor to the evolution of the healthcare marketplace—toward a more market-oriented focus—has been the application of the antitrust laws to the healthcare industry.[16] Elsewhere, I have observed that "[a]ntitrust law is the virtual engine of the market paradigm."[17] Antitrust focuses on the promotion of competition and evaluates conduct according to considerations of economic efficiency and consumer welfare. Because so much of policy in the healthcare arena has been driven by equitable concerns regarding access to quality medical care, the enforcement of antitrust in the healthcare industry raises an inevitable tension. Market efficiency may result in the more appropriate use of resources, and improved competition and efficiency may result in economies that benefit consumers who might otherwise not be able to afford those services. However, even with an efficient system, there will be persons whose income is just too low to pay for medical care.

Traditionally, the healthcare system has used cross subsidies to achieve "worthy purposes," such as the financing of services for those without the resources to pay for medical care on their own.[18] The funds for this cross subsidization have stemmed from the receipt (typically by hospitals) of supra-competitive returns in some areas; those supra-competitive returns reflect the ability of hospitals to exert a form of monopoly control in certain market niches, allowing for the receipt of revenues beyond a competitive return. By focusing on promoting competition and economic efficiency, and by barring anticompetitive conduct that leads to the earning of supra-competitive returns, antitrust laws constrain the ability of providers and provider institutions to achieve supra-competitive returns. That, in turn, compromises the ability of

---

13. *See* Clark C. Hauighurst & James F. Blumstein, *Coping with Quality/Cost Trade-Offs in Medical Care: The Role of PSROs*, 70 NW. U. L. REV. 6, 9-11 (1975).

14. *See* Randall R. Bovbjerg, *Competition Versus Regulation in Medical Care: An Overdrawn Dichotomy*, 34 VAND. L. REV. 965, 1001 (1981).

15. For a general discussion of these issues, see James F. Blumstein, *Health Care Reform and Competing Visions of Medical Care: Antitrust and State Provider Cooperation Legislation*, 79 CORNELL L. REV. 1459, 1463-86 (1994).

16. Thomas E. Kauper, *The Role of Quality of Health Care Considerations in Antitrust Analysis*, 51 LAW & CONTEMP. PROBS., Spring 1988, at 273.

17. Blumstein, *supra* note 15, at 1482.

18. David M. Frankford, *Creating and Dividing the Fruits of Collective Economic Activity: Referrals Among Health Care Providers*, 89 COLUM. L. REV. 1861, 1938 (1989).

healthcare institutions such as hospitals to cross subsidize.[19] The competing away of supra-competitive returns is a natural result of the introduction of competition; also, because antitrust circumscribes anticompetitive collusive or monopolistic conduct, it limits the ability of providers and provider institutions to restore their ability to cross subsidize by earning supra-competitive returns.[20] Many of the steps necessary to achieve those supra-competitive returns will subject an institution to antitrust enforcement scrutiny.

I. THE SUBSTANTIVE AND SYMBOLIC IMPORTANCE OF ANTITRUST

Antitrust law has both substantive and symbolic importance. "In very fundamental ways, application of antitrust principles to the medical care arena transforms thinking about certain issues." [21]

Application of the antitrust laws to healthcare alters the way that participants think about the services being provided and received—about the very nature of the healthcare enterprise. Thus, by changing the culture and the climate of the entire healthcare arena, the antitrust law is important in symbolic terms.

For example, under the health planning umbrella,[22] policymakers encouraged healthcare institutional managers to rationalize the "system" of healthcare delivery. Healthcare was considered a "system," with all that term connotes. Policymakers and planners would spend a good bit of time considering what institutional design structure would best achieve efficiencies and provide patients with accessible, high-quality services.[23]

Use of the term "system" suggests a social services delivery model. In that context, the term "non-system" is a pejorative, connoting that there should be an organized system, but that there is not one. Consistent with the health planning approach, much criticism was leveled at the American healthcare delivery "system" because it was insufficiently organized or inadequately structured.

Application of the antitrust laws to the healthcare arena makes it clear that issues involving "trade or commerce" are at stake.[24] That is, healthcare is an

---

19. *See* Blumstein, *supra* note 15, at 1500-01.

20. *Id.* at 1482-86.

21. *Id.* at 1482.

22. *See generally* James F. Blumstein, *Effective Health Planning in a Competitive Environment*, *in* COST, QUALITY, AND ACCESS IN HEALTH CARE: NEW ROLES FOR HEALTH PLANNING IN A COMPETITIVE ENVIRONMENT 21 (Frank A. Sloan et al. eds., 1988); Randall R. Bovbjerg, *New Directions for Health Planning*, *in* COST, QUALITY, AND ACCESS IN HEALTH CARE, *supra*, at 206.

23. *See generally* DAVID D. RUTSTEIN, BLUEPRINT FOR MEDICAL CARE (1974) (considering how to best organize medical care under a putative national health insurance scheme).

24. Although the antitrust laws have been applied in some situations involving healthcare for a half-century, *see* AMA v. United States, 317 U.S. 519, 528 (1943), it was the *Goldfarb* decision in 1975 that definitively determined that there was no generalized antitrust exemption for the "learned professions." Goldfarb v. Virginia State Bar, 421 U.S. 773, 787 (1975). In *Goldfarb*, the Supreme Court recognized that the practice of law had business dimensions and was "trade or

"industry" to be policed through antitrust enforcement against anticompetitive conduct as are other economic sectors. In an "industry," principles of economics have application. Economic concepts such as supply and demand, incentives, and trade-offs become important terms for analysis and consideration. The "non-system" terminology has limited applicability in the context of an economic market, which is typically driven by decentralized decisions of individual households and firms.[25]

Viewing healthcare as an industry ("trade or commerce") rather than purely as a social services delivery system is truly transformative in terms of the culture of the actors within the industry. The very same activity that is considered appropriate or constructive under one behavioral framework (*e.g.*, the professional/planning model) is viewed as harmful and even illegal under another (*e.g.*, the competitive/antitrust model).

Under health planning, as influenced by the professional paradigm, institutional providers such as hospitals were encouraged to act collectively to eliminate "wasteful duplication."[26] Cooperation and coordination were seen as socially appropriate tools for rationalizing a "system" by participants in that system in pursuit of the common good. From an antitrust perspective, however, such conduct between or among competitors is far from the wholesome activity envisioned by its health planning proponents. Whereas "cooperation" or "coordination" seem like good things, antitrust enforcers are likely to see them in less glowing terms. "Cooperation" or "coordination" becomes the much more perverse "conspiracy;" the purported elimination of "wasteful duplication" looks like an illegal restraint of trade among competitors—a form of territorial market division that is so destructive to competition that the Supreme Court has labeled such restraints as per se illegal violations of the antitrust laws.[27]

---

commerce" as required for coverage under the Sherman Anti-Trust Act, ch. 647, 26 Stat. 209 (1890) (current version at 15 U.S.C. § 1 (1994)). *Goldfarb*, 421 U.S. at 787-88. Following *Goldfarb*, the Court clearly indicated that the antitrust laws applied generically to professionals. *See* National Soc'y of Prof'l Eng'rs v. United States, 435 U.S. 679, 695 (1978). Despite the existence of a federal regime of health planning at the time (since repealed), the Court also made clear that the antitrust laws applied to the healthcare arena. *See* National Gerimedical Hosp. & Gerontology Ctr. v. Blue Cross, 452 U.S. 378, 391 (1981). Thus, Congress was forced to enact a limited exemption from antitrust coverage for hospital peer review activities, which were deemed essential for the maintenance of quality assurance procedures. *See* Health Care Quality Improvement Act of 1986, 42 U.S.C. §§ 11101-11152. *See generally* James F. Blumstein & Frank A. Sloan, *Antitrust and Hospital Peer Review*, LAW & CONTEMP. PROBS., Spring 1988, at 7. It is now clear that the antitrust laws apply to conduct by professionals in the healthcare arena. *See* Summit Health, Ltd. v. Pinhas, 500 U.S. 322, 329-30 (1991); Patrick v. Burget, 486 U.S. 94, 101-05 (1988).

25. James F. Blumstein, *Rationing Medical Resources: A Constitutional, Legal and Political Analysis*, 59 TEX. L. REV. 1345-1400 (1981).

26. *See* Clark C. Hauighurst, *Regulation of Health Facilities and Services by "Certificate of Need,"* 59 VA. L. REV. 1143, 1148-51 (1973).

27. *See* United States v. Topco Assocs., Inc., 405 U.S. 596, 608-11 (1972). *Cf.* Hauighurst,

Clearly, then, the application of the antitrust laws to the healthcare industry has a critical symbolic function in altering the perception of how to think about the healthcare enterprise—from a social services delivery system to an industry in the economic marketplace. Accompanying the symbolic importance of antitrust is a critical substantive dimension as well.

One of the hallmarks of a profession is its sense of self-regulation. Professionals assume certain power by substituting their judgment on critical technical and economic matters for that of consumers.[28] The social quid pro quo for this empowerment is an understanding that professionals should not act in their own economic self interest at the expense of consumers. In theory, at least, self-regulation by the profession itself substitutes for the discipline of the marketplace, which normally determines price, quality, and level of services provided.

The frame of reference for professionals is the peer group of professionals. Members of the peer group establish ethical codes of conduct that govern standards of behavior within a profession. Professional peers set standards of quality[29] and, through various mechanisms of peer review, enforce standards of conduct within the profession. In short, it is entirely within the professional tradition for groups of peers to act collectively to establish and enforce professional norms. In some situations, this custom of joint conduct has been used to establish or retain professional hegemony or to resist threats against perceived inroads on providers' autonomy or financial well-being.[30]

For example, physicians have acted collectively against unwanted competition,[31] have engaged in group boycotts to resist fee pressures,[32] and have resisted alternative payment methods to combat the growth and development of prepaid plans such as health maintenance organizations (HMOs),[33] and dentists have acted in concert to resist cost-containment efforts by an insurer that threatened provider incomes and autonomy.[34] Antitrust norms call this type of collective behavior into question.

Further, the professional instinct for collective conduct to assert and to

---

supra note 26, at 1149 ("[M]any of the activities undertaken in the name of planning were indistinguishable from such typical cartel practices as output restriction (collective determination of the bed supply) and market division (allocation of areas of responsibility both geographically and by activity.)"). *See also* Blumstein & Sloan, *supra* note 24, at 8-9 (discussing this dissonance).

28. *See* Greaney, *supra* note 2, at 633-35.

29. This is particularly true in the field of medicine, where liability standards for professional negligence are determined by reference to professional norms of conduct. *See* Blumstein & Sloan, *supra* note 24 and accompanying text.

30. *See, e.g.*, Patrick v. Burget, 486 U.S. 94 (1988); FTC v. Indiana Fed'n of Dentists, 476 U.S. 447 (1986); AMA v. United States, 317 U.S. 519 (1943); Nurse Midwifery Assocs. v. Hibbett, 918 F.2d 605 (6th Cir. 1990), *modified*, 927 F.2d 904 (6th Cir.), *cert. denied*, 502 U.S. 952 (1991).

31. *See, e.g.*, *Patrick*, 486 U.S. at 94.

32. *See, e.g.*, *In re* Michigan State Med. Soc'y, 101 F.T.C. 191 (1983).

33. *See, e.g.*, *AMA*, 317 U.S. at 519.

34. *See, e.g.*, *Indiana Fed'n of Dentists*, 476 U.S. at 447.

enforce collectively determined norms of conduct is subject to challenge when confronted by the antitrust laws' norm against collective conduct to discipline the behavior of competitors. In the absence of Congressionally-conferred specific legislative exemptions,[35] courts applying the antitrust laws have expressed reluctance to weigh procompetitive virtues against other competing policy objectives. Thus, in an antitrust case, courts will routinely balance procompetitive against anticompetitive aspects of a restraint or a set of restraints. What antitrust courts tend to eschew, however, is the temptation to balance procompetitive values embodied in the antitrust laws with policy objectives other than those associated with competition.[36] If goals related to other policy values are to trump the procompetitive virtues of the marketplace protected by the antitrust laws, the courts have found that only Congress can carve out such an exception to antitrust coverage.[37]

This fundamental rule of antitrust law has been applied even to bar a group boycott aimed at disciplining illegal activity.[38] This, in many ways, is the essence of professional self regulation, yet the antitrust courts have typically been unwilling to allow that kind of concerted action because of fears of harm to the competitive process. The Supreme Court has applied this doctrine in the professional context, barring as illegally anticompetitive collective professional activity that assertedly promotes values other than competition (such as quality of services, a hallmark of professional self-regulation).[39] Indeed, in condemning challenges to antitrust enforcement actions based on the asserted worthiness of the pursuit of other, noncompetitive objectives—the "worthy purpose" defense—the Court has labeled such purported defense theories a "frontal assault" on the core values and policies of the antitrust laws themselves.[40]

Thus, from a substantive viewpoint, the antitrust laws are important in part because they eliminate non-efficiency-based criteria from consideration in deciding the legality of collective conduct. This comes into conflict with traditional professional norms, which recognize values other than competition as justifications for collective conduct.[41] Antitrust enforcement, therefore,

---

35. *See Patrick*, 486 U.S. at 105 & 106 n.8.

36. *See, e.g.*, National Soc'y of Prof'l Eng'rs v. United States, 435 U.S. 679 (1978).

37. *See Patrick*, 486 U.S. at 105 & 106 n.8. *See generally* Blumstein & Sloan, *supra* note 24, at 28-32; Greaney, *supra* note 2; Kauper, *supra* note 16.

38. Fashion Originators' Guild of Am., Inc. v. FTC, 312 U.S. 457, 464 (1941).

39. *See, e.g.*, *Indiana Fed'n of Dentists*, 476 U.S. at 447; *National Soc'y of Prof'l Eng'rs*, 435 U.S. at 679.

40. *See Nat'l Soc'y of Prof'l Eng'rs*, 435 U.S. at 680. *But see* United States v. Brown Univ., 5 F.3d 658, 661 (3d Cir. 1993); Wilk v. AMA, 719 F.2d 207, 222 (7th Cir. 1983).

41. The American Hospital Association (AHA), for example, advocated adoption of state laws that, making use of the antitrust exemption for state action under *Parker v. Brown*, 317 U.S. 341 (1943), would authorize collaborative activities among healthcare providers that might otherwise violate the federal antitrust laws. The rationale, in part, for that AHA initiative was dissatisfaction with antitrust doctrine that competition as a value cannot be offset, in an antitrust analysis, by other values such as improved access to or quality of medical care. *See* Fredric J. Entin

confronts the professional ideal that professionals control basic economic decisions regarding price, quality, and level of services provided and that professionals can act collectively to enforce those professional norms. Antitrust enforcers are notoriously skeptical of these professional claims; and antitrust enforcement in the healthcare arena compels professionals and providers to give more attention to traditional economic considerations of balancing quality and cost. Sensitivity on the part of providers to concerns of payors and consumers with regard to access, quality, and cost is a likely consequence. Accommodation to consumer desires to share in their medical care decision making may also be a result of enhanced attention to procompetitive considerations. From the perspective of market reform, it is important symbolically and substantively to maintain the role of the antitrust enforcement which has helped change the way that policymakers and market participants think about medical care.

II. THE SIGNIFICANCE OF EMPIRICAL ISSUES IN ANTITRUST ENFORCEMENT

Given the normative importance of the application of the antitrust laws to the healthcare environment, one could reasonably predict that there would be conflict regarding empirical issues that undergird the application of the antitrust laws to any industry. For example, conventional antitrust doctrine requires the use of empirical analysis to determine such issues as the existence or nonexistence of market power—the ability of an economic actor or set of actors to influence price or quantity in a market. Similarly, empirical evidence will often be sought regarding the competitive effects of various restraints in a particular market. That inquiry is necessary in a traditional rule-of-reason analysis to determine whether, on balance, a particular restraint is procompetitive or anticompetitive.[42]

To determine the issue of market power, an analyst must define a geographic market and a product market. After all, it is impossible to determine whether an economic actor or set of actors exerts market power without knowing what the market is. The determination of the market, in turn, is composed of a number of empirical considerations regarding geography and definition of the product involved. For example, in determining what the geographic market is, analysts must determine how far consumers will travel to consume a specific type of medical service when facing different levels of price or perceived quality.[43] This issue is nearly always an essential part of a rule-of-reason antitrust analysis and turns on empirical evidence specifically related to a particular market and set of services. There may be a difference in the scope of competition in different types of geographic markets depending on the nature of the medical services involved. For example, it is often suggested that the geographic market for primary care services is more narrowly circumscribed than the market for more specialized

---

et al., *Hospital Collaboration: The Need For an Appropriate Antitrust Policy*, 29 WAKE FOREST L. REV. 107, 127-28, 134 (1994).

42. *National Soc'y of Prof'l Eng'rs*, 435 U.S. at 692.

43. *See, e.g.*, Michael A. Morrisey et al., *Defining Geographic Markets for Hospital Care*, LAW & CONTEMP. PROBS., Spring 1988, at 165.

tertiary care services.[44]

In defining the appropriate product market, analysts must determine what products or services compete with each other. That is, what products or services can be substituted for others. This is an empirical question that focuses on what economists call the cross elasticities of demand—what products or services can substitute for others and at what prices. For example, at what point do primary care physicians compete with specialists? Presumably, there is some continuum of quality and price at which payors and consumers will substitute primary care services for specialty services.

One important note of caution is necessary in evaluating empirical evidence in the context of antitrust policy. Empirical evidence is not developed in an institutional vacuum. Behavior is shaped by structural incentives, so observed consequences must be evaluated in the context of the institutional milieu in which the empirical observation is made. Thus, in considering the consequences of increased competition in a particular market, one must be aware of the institutional constraints on how competition is channeled before drawing policy conclusions. This insight will be important in the ensuing discussion of how empirical evidence should be evaluated in an antitrust policy context.

A. General

In definitively applying the antitrust laws to the professions, the Supreme Court recognized that the empirical effects of collective conduct in the context of the professions (such as medicine) might differ from that in other economic sectors.[45] The Court seemingly has rejected the "worthy purpose" defense in the context of the professions.[46] An example of this is disallowing the balancing of procompetitive values against non-competitive policy objectives designed to achieve worthy policy objectives through anticompetitive practices. The Court has steadfastly acknowledged, however, that not all doctrinal conclusions developed in other economic sectors will automatically transfer to the realm of the professions.

Thus, the Court has not been willing, as a matter of routine, to apply rules developed in other economic sectors to the professional context, preferring to allow defendants to proffer ostensible procompetitive justifications for independent judicial evaluation.[47] This is perhaps best illustrated in the area of

---

44. *See, e.g.*, Blue Cross & Blue Shield United v. Marshfield Clinic, 65 F.3d 1406 (7th Cir. 1995); United States v. Long Island Jewish Med. Ctr., No. CV973412 (ADS), 1997 WL 662731 at *50 (E.D.N.Y. Oct. 23, 1997) (finding as a fact that different geographic markets existed for primary/secondary hospital services and for tertiary care services).

45. *See* Goldfarb v. Virginia State Bar, 421 U.S. 773, 788-89 n.17 (1975).

46. *See* FTC v. Indiana Fed'n of Dentists, 476 U.S. 447, 463 (1986); *National Soc'y of Prof'l Eng'rs*, 435 U.S. at 695-96; Blumstein & Sloan, *supra* note 24, at 28-32.

47. Some rules, such as the per se ban on price fixing, have been applied in the medical care and professional context. *See* FTC v. Superior Ct. Trial Lawyers, 493 U.S. 411, 432 (1990); Arizona v. Maricopa County Med. Soc'y, 457 U.S. 332, 348 (1982).

the antitrust rules of per se invalidity under section 1 of the Sherman Act.[48]

Section 1 requires a threshold finding that there is concerted action by more than one party with the capacity to agree or conspire.[49] Once that threshold requirement is satisfied, the most fundamental practical analytical issue becomes which standard of antitrust to apply: the rule of per se invalidity or the rule of reason. Under traditional antitrust doctrine, the per se approach means that in carefully delineated situations an antitrust complainant need only establish the agreement itself to prevail in the litigation. Under a per se analysis, courts do not inquire elaborately into the precise nature or scope of the purported harm to competition or into the possible business justifications for use of the challenged practice.[50] Per se rules have developed from judicial experience, where courts conclude that a particular type of concerted conduct has a "pernicious effect on competition and lack[s] . . . any redeeming virtue."[51] Under such circumstances, use of per se rules is efficient, saving time and lowering complainant's cost of putting forward an expensive antitrust action. Because of their open-and-shut character, per se rules are carefully targeted, and the courts have been reluctant to be expansive in interpreting their scope.[52]

Practices that are subject to review under section 1 but that do not warrant per se treatment are anlyzed under the rule of reason. The test of legality under the rule of reason is "whether the restraint imposed is such as merely regulates and perhaps thereby promotes competition or whether it is such as may suppress or even destroy competition."[53] A court must determine the "competitive significance of the restraint" and evaluate the "facts peculiar to the business, the history of the restraint, and the reasons why it was imposed."[54] The bottom line question a court must decide under the rule of reason is whether, on balance, a "restrictive practice should be prohibited as imposing an unreasonable restraint on competition."[55]

Application of the rule of per se invalidity does not allow a defendant to justify its conduct at all; proof of an agreement or a conspiracy that triggers a per se violation pretermits consideration of justifications. In the professional context, the Court normally has been willing to hear and consider purported

---

48. 15 U.S.C. § 1 (1994).

49. *See* Copperweld Corp. v. Independence Tube Corp., 467 U.S. 752, 768 (1984). *See* generally Blumstein & Sloan, *supra* note 24, at 39-53.

50. *See* United States v. Topco Assocs., Inc., 405 U.S. 596, 607 (1972); *see also National Soc'y of Prof'l Eng'rs*, 435 U.S. at 692.

51. Northern Pac. Ry. Co. v. United States, 356 U.S. 1, 5 (1958).

52. *See* Northwest Wholesale Stationers, Inc. v. Pacific Stationery & Printing Co., 472 U.S. 284, 294-95 (1985).

53. Board of Trade v. United States, 246 U.S. 231, 238 (1918). "Under the rule of reason, the court must determine whether the consequences of the contested concerted conduct constitute contraindicated constraints on competition." Blumstein & Sloan, *supra* note 24, at 54.

54. *National Soc'y of Prof'l Eng'rs*, 435 U.S. at 692.

55. Continental T.V., Inc. v. GTE Sylvania Inc., 433 U.S. 36, 49 (1977).

procompetitive rationales for restraints.[56] Indeed, the professional context has spawned a doctrinal accommodation: a third, intermediate level of scrutiny, in which plaintiffs need only demonstrate what appears to be a suspiciously naked restraint. While the defendants in these cases are afforded an opportunity to defend their collective conduct on the basis of a procompetitive rationale, the burden of justification rests with them, and plaintiffs need not present detailed evidence of market structure or particularized harm to competition.[57] Thus, professional defendants are usually afforded an opportunity to present procompetitive justifications for contested restraints, and that provides an opportunity for professionals to draw on theoretical or empirical evidence to rebut a claim that the challenged conduct should be held illegal.

In short, the federal antitrust laws, as applied to the healthcare arena, contemplate a careful evaluation of the competitive impact of various restraints on competition.[58] While the Court has warned against "fashioning a broad exemption under the Rule of Reason for learned professions" on the theory that "competition itself is unreasonable,"[59] it has also acknowledged that "by their nature, professional services may differ significantly from other business services" in their competitive impact.[60] Thus, "the nature of the competition in such services may vary."[61] A restraint on competition cannot be justified "on the basis of the potential threat that competition poses to the public safety" because that would be "nothing less than a frontal assault on the basic policy of the Sherman Act."[62] But antitrust enforcement must be sensitive to the particular context in which a restraint arises so that its competitive impact is fully understood. In sum, the federal antitrust laws "provide a great degree of flexibility for private collaborative efforts aimed at achieving more efficient and less costly delivery of health care services."[63] In the healthcare arena, empirical analysis will ultimately determine whether specific restraints peculiar to the healthcare industry will be vindicated as procompetitive or struck down as unwarranted restraints on competition.[64]

---

56. *See* FTC v. Indiana Fed'n of Dentists, 476 U.S. 447 (1986); *National Soc'y of Prof'l Eng'rs*, 435 U.S. at 693. While this is generally true, there are exceptions, notably the holding that the rule of per se invalidity applies to price fixing agreements among professionals. *See* Arizona v. Maricopa County Med. Soc'y, 457 U.S. 332 (1982); *see also* FTC v. Superior Ct. Trial Lawyers, 493 U.S. 411, 432 (1990).

57. *See Indiana Fed'n of Dentists*, 476 U.S. at 459; *see also* United States v. Brown Univ., 5 F.3d 658, 674 (3d Cir. 1993) (explaining intermediate standard of antitrust scrutiny).

58. *See* Goldfarb v. Virginia State Bar, 421 U.S. 773, 788-89 n.17 (1975).

59. *National Soc'y of Prof'l Eng'rs*, 435 U.S. at 696.

60. *Id.*

61. *Id.*

62. *Id.* at 695.

63. David L. Meyer & Charles F. Rule, *Health Care Collaboration Does Not Require Substantive Antitrust Reform*, 29 WAKE FOREST L. REV. 169, 171 (1994).

64. *See Goldfarb*, 421 U.S. at 788-89 n.17. In recognition of the need for developing antitrust enforcement policy in the specific context of the healthcare industry, the Department of

B. The Case of Hospital Cooperation Laws

Many states have enacted laws designed to insulate conduct of hospitals or health care providers from scrutiny under the antitrust laws.[65] Although the rationales for these laws have been called into question,[66] their proponents claim that they are needed, in part, to promote economic efficiency.[67] While part of the rationale for these laws seems to be based on the desire to allow health care providers to cross subsidize to achieve worthy health policy objectives and to avoid the impetus toward efficiency of antitrust enforcement,[68] a substantial contributor to the enactment of these laws was an expressed concern with the ability of federal antitrust enforcement agencies to understand the economic realities of the healthcare marketplace.[69]

These so-called hospital cooperation laws are designed to make use of the *Parker v. Brown*[70] state-action immunity doctrine. Under general constitutional principles (the Supremacy Clause[71]), federal law supersedes state law that comes in conflict with it. In the realm of antitrust, however, the Supreme Court has construed the federal antitrust law to embrace a principle of federalism[72] which

---

Justice and the Federal Trade Commission have adopted enforcement safety zones to provide guidance to industry participants of how the antitrust enforcement agencies interpret the antitrust laws in their application to the healthcare industry. U.S. Department of Justice & Federal Trade Commission, *Statements of Antitrust Enforcement Policy in Health Care* (Aug. 28, 1996) <http://www.ftc.gov/reports/hlth3s.htm>. The safety zones have been promulgated not as exceptions to, but as enforcement guidelines within, the framework of general antitrust policies. This point was made clear by a statement of Senator Howard Metzenbaum issued when the original safety zones were announced:

> We're going to solve a problem in the antitrust field without changing one word, one comma, or one semicolon of the antitrust laws. . . . We are here today to clear up confusion among doctors and hospitals about how these laws apply to them. We want to end their uncertainty. . . . These policy guidelines are proof positive that we can make our laws work to accommodate business when their concerns have logic and merit.

Statement by Senator Howard M. Metzenbaum (Sept. 15, 1993), *Press Conference with U.S. Att'y Gen. Janet Reno, First Lady Hillary Clinton, Janet Steiger, Chair, Federal Trade Commission*, Fed. News Service, Sept. 15, 1993, *available in* LEXIS, News Library, ARCNWS file.

65. For a list of these states as of October 1995, see James F. Blumstein, *Assessing Hospital Cooperation Laws*, 8 LOY. CONSUMER L. REP. 248, 253-54 (1996) (Table 1).

66. *See* Blumstein, *supra* note 15, at 1493-1501; Meyer & Rule, *supra* note 63, at 176-82.

67. *See generally* Entin et al., *supra* note 41.

68. *See* Blumstein, *supra* note 15, at 1498.

69. *Compare* Meyer & Rule, *supra* note 63, *with* Entin et al., *supra* note 41.

70. 317 U.S. 341 (1943).

71. U.S. CONST. art. VI ("This Constitution, and the Laws of the United States which shall be made in Pursuance thereof . . . shall be the supreme Law of the Land . . . any Thing in the Constitution or Laws of any State to the Contrary notwithstanding.").

72. For a discussion of this principle of statutorily-mandated federalism, see James F.

authorizes states to overcome the effect of federal antitrust enforcement by substituting a regime of state regulation for competition.[73] Thus, federal antitrust law does not apply "to anticompetitive restraints imposed by the States 'as an act of government,'"[74] and is "subject to supersession by state regulatory programs,"[75] provided that the state must clearly articulate its policy to substitute regulation for competition and must actively supervise implementation of that policy.[76]

These hospital cooperation laws assume that "in the health care marketplace regulation may be preferable to competition in some circumstances."[77] To the extent that these laws are premised on the use of regulation to promote considerations of economic efficiency, they demonstrate a resistance to the application of antitrust as a vehicle for the enforcement of competition and thus economic efficiency. These laws reflect a "resistance to the shift in paradigms" from the professional to the market-oriented approach and assume that "features of the health care marketplace that impede the proper functioning of the market will remain in place."[78]

Importantly, advocacy of these laws was based on empirical evidence that competition among health care providers can result in the encouragement of wasteful practices.[79] This brings home a critical point regarding the use of empirical evidence in the analysis and evaluation of antitrust enforcement in the healthcare arena—empirical evidence will be shaped by the institutional structure of the industry. If avenues for competition are circumscribed, then competition can result in certain perverse or cost-escalating outcomes. But the policy implication of these studies must be evaluated in the context of possible institutional reform or redesign, which would change the nature and effect of competition. Similarly, antitrust enforcement authorities must be aware of the shifting nature of competition in the healthcare marketplace so that enforcement

---

Blumstein, *Federalism and Civil Rights: Complementary and Competing Paradigms*, 47 VAND. L. REV. 1251, 1294-1300 (1994). *See generally* James F. Blumstein & Terry Calvani, *State Action as a Shield and a Sword in a Medical Services Antitrust Context*: Parker v. Brown *in Constitutional Perspective*, 1978 DUKE L.J. 389, 395-97, 400-03, 414-31.

73. *See* Frank H. Easterbrook, *Antitrust and the Economics of Federalism*, 26 J.L. & ECON. 23, 25 (1983) (describing *Parker* as a form of "inverse preemption"); *but see* Einer R. Elhauge, *The Scope of Antitrust Process*, 104 HARV. L. REV. 667, 717-29 (1991) (critiquing that view of the state-action doctrine).

74. City of Columbia v. Omni Outdoor Adver., Inc., 499 U.S. 365, 370 (1991) (Stevens, J., dissenting).

75. FTC v. Ticor Title Ins. Co., 504 U.S. 621, 632-33 (1992).

76. California Retail Liquor Dealers Ass'n v. Midcal Aluminum, Inc., 445 U.S. 97, 105 (1980).

77. Blumstein, *supra* note 15, at 1490.

78. *Id.* at 1494.

79. *See* James C. Robinson & Harold S. Luft, *The Impact of Hospital Market Structure on Patient Volume, Average Length of Stay, and the Cost of Care*, 4 J. HEALTH ECON. 333, 353-54 (1985).

policies in fact, as well as in theory, promote competition rather than promote outdated rules of thumb that can, under the changed circumstances, stand in the way of appropriate procompetitive conduct.[80]

In the debate surrounding adoption of the hospital cooperation legislation, advocates of the legislation contended that the economic structure of the healthcare industry meant that price competition was not suitable.[81] Evaluation of that argument requires some consideration of the historical structure of the healthcare market.

Traditionally, physicians have been the most influential participants in the healthcare market.[82] This influence has stemmed from the expertise and knowledge physicians have obtained from their specialized training[83] and from their control of patients and patient referrals.[84] The professional dominance model[85] has both resulted from[86] and reinforced this asymmetry of information.[87] With professional dominance, patients tend to rely on the recommendations of their physician regarding referrals.

This traditional ability of physicians to channel patients has meant that hospitals have been dependent on physicians to admit patients to their facilities.[88] Understandably, in such circumstances and with the prevalence of third-party insurance for hospital stays offsetting most out-of-pocket patient expenses,[89]

---

80. *See* Clark C. Hauighurst, *Are the Antitrust Agencies Overregulating Physician Networks?*, 8 LOY. CONSUMER L. REP. 78 (1995-96).

81. *See* Entin et al., *supra* note 41, at 122-38.

82. *See* STARR, *supra* note 2, at 226-27; Mark A. Hall, *Institutional Control of Physician Behavior: Legal Barriers to Health Care Cost Containment*, 137 U. PA. L. REV. 431, 445-47 (1988).

83. *See* J. Michael Woolley & H.E. Frech, III, *How Hospitals Compete: A Review of the Literature*, 2 U. FLA. J.L. & PUB. POL'Y 57, 59-60 (1988-89).

84. *See* Harold S. Luft et al., *The Role of Specialized Clinical Services in Competition Among Hospitals*, 23 INQUIRY 83 (1986). *See generally* Blumstein & Sloan, *supra* note 24, at 17 ("Since doctors have traditionally referred patients to hospitals, they have controlled the hospitals' clientele. That power . . . has given physicians . . . considerable leverage over hospitals.").

85. *See* Blumstein, *supra* note 15, at 1463-64.

86. *See* STARR, *supra* note 2, at 226-27. Starr has argued that the dominance of professionals has perpetuated the imbalance in information available to patients, and, thereby, has perpetuated professional power vis a vis patients. In Starr's account, professionalism may in part be a cause, not exclusively a response, to market failure (the asymmetry of information between physician and patient).

87. *See* Arrow, *supra* note 2, at 947-49 (arguing that the professional paradigm is a response to the market failure in the medical care marketplace—the unpredictable nature of the need for medical care and the asymmetry of information (the knowledge of the physician and the ignorance of the consumer)). *But see* STARR, *supra* note 2, at 226-27 (noting that uncertainty and consumer ignorance may be promoted by the professional paradigm, thereby perpetuating the empowerment of professionals in medical care decisionmaking).

88. *See* Luft et al., *supra* note 84.

89. *See* Woolley & Frech, *supra* note 83, at 60-61.

competition among hospitals focused on attracting referrals of patients by physicians. In that type of competitive environment, still prevalent in many parts of the United States, emphasis among competing hospitals is on appealing to the physician-referrers; where third-party reimbursement is relatively automatic and cost-based,[90] neither the hospital nor the physician has much of an incentive to be particularly responsive to considerations of cost.

In an industry with such structural characteristics, it is surely no surprise to learn that increased competition would be associated with higher prices[91]—the so-called "medical arms race," where purchases of expensive equipment by one institution led to similar purchases by competitor institutions without regard for cost effectiveness and without regard to constraints of cost.[92] Since physicians controlled patient flow to hospitals through control of patient referrals, hospitals often competed for physician affiliations by providing expensive specialized clinical services[93] with the attendant escalation in capital expenditure, increases in overhead and the concomitant increase in operating costs.[94] In such an environment, empirical evidence suggested that hospitals in more competitive markets experienced higher costs.[95]

Although empirical evidence may have supported the "medical arms race" hypothesis in the presence of a particular market structure that channeled competition in certain ways, subsequent empirical research strongly suggests that, when the institutional structure of the healthcare marketplace changed, the impact of competition also changed. That is, when institutional and legal change altered the structure of the healthcare market[96] so that it reflected the kinds of economic incentives present in other markets, it seems that participants in the healthcare market behaved much like participants in other markets.[97]

For example, when selective contracting was broadly introduced into the

---

90. In a cost-based mode of hospital reimbursement, greater competition may be associated with higher rather than lower costs. *See* Robinson & Luft, *supra* note 79, at 353-54.

91. *See, e.g.*, James C. Robinson & Harold S. Luft, *Competition and the Cost of Hospital Care, 1972 to 1982*, 257 JAMA 3241, 3244 (1987); Jack Zwanziger & Glenn A. Melnick, *The Effects of Hospital Competition and the Medicare PPS Program on Hospital Cost Behavior in California*, 7 J. HEALTH ECON. 301, 305 (1988). For a more generalized discussion of the relationship between the nature of competition and the containment of costs, see Thomas L Greaney, *Managed Competition, Integrated Delivery Systems and Antitrust*, 79 CORNELL L. REV. 1507, 1513-14 (1994).

92. *See* Luft et al., *supra* note 84, at 92.

93. *See id.* at 83; Hall, *supra* note 82, at 506.

94. *See* Luft et al., *supra* note 84, at 93; Robinson & Luft, *supra* note 91, at 3241.

95. *See* Robinson & Luft, *supra* note 79, at 342 ("Hospitals in monopolistic positions within their local area produce[d] their services at significantly lower costs than hospitals in more competitive environments."). *See generally* United States v. Carilion Health Sys., 707 F. Supp. 840, 846 (W.D. Va. 1989).

96. The data for these subsequent studies derived largely from California.

97. *See* discussion *supra* notes 1-20 and accompanying text.

California market,[98] allowing private health plans and Medi-Cal[99] to channel patients to selected providers in exchange for price and other concessions, price competition was introduced into the California healthcare market.[100] With the broader introduction of cost-conscious payers into the healthcare market, incentives shifted, and, as a result, price competition as well as quality competition began to emerge.[101]

The altered payment policies in California in the early 1980s, and the resultant shift in economic incentives for participants in the healthcare marketplace substantially reduced the rate of increase in total hospital costs and revenues, and caused a shift to less expensive outpatient services.[102] With strong incentives to reduce costs,[103] hospitals experienced a lower growth rate in costs in the 1983-85 period than the 1980-82 period in all categories except for outpatient services.[104] For hospitals in highly competitive areas, total inpatient costs adjusted for inflation declined by 11.3% while remaining flat in low-competition markets.[105] From 1983 to 1988, high HMO penetration stimulated more price competitive behavior on the part of traditional health insurers. When those insurers were allowed to negotiate and contract with hospitals for discounts, they did so, with a reduction in costs.[106]

The empirical evidence from California strongly suggests that in competitive hospital markets, when appropriately structured, the traditional economic

---

98. *See* Glenn A. Melnick & Jack Zwanziger, *Hospital Behavior under Competition and Cost-Containment Policies: The California Experience, 1980 to 1985*, 260 JAMA 2669, 2669 (1988); James C. Robinson, *HMO Market Penetration and Hospital Cost Inflation in California*, 266 JAMA 2719, 2719 (1991).

99. Medi-Cal is California's Medicaid program.

100. This allowed for bargaining on the part of private health plans and Medi-Cal with providers. *See* David Dranove & William D. White, *Recent Theory and Evidence on Competition in Hospital Markets*, 3 J. ECON. & MGMT. STRATEGY 169, 193-94 (1994); Melnick & Zwanziger, *supra* note 98, at 2669; James C. Robinson & Harold S. Luft, *Competition, Regulation, and Hospital Costs, 1982 to 1986*, 260 JAMA 2676, 2676 (1988); Robinson, *supra* note 98, at 2719; Zwanziger & Melnick, *supra* note 91, at 316-17; Jack Zwanziger et al., *Hospitals and Antitrust: Defining Markets, Settings Standards*, 19 J. HEALTH POL., POL'Y & L. 423, 424 (1994).

101. *See* Melnick & Zwanziger, *supra* note 98, at 2675; Robinson, *supra* note 98, at 2723; David Dranove et al., *Price and Concentration in Hospital Markets: The Switch from Patient-Driven to Payer-Driven Competition*, 36 J.L. & ECON. 179, 180-81 (1993). The rate of increase in inpatient costs adjusted for inflation increased at an average rate of almost 5% in 1980-82 and decreased by almost 2% in the 1983-85 periods, after the introduction of selective contracting and the introduction of price negotiation. Melnick & Zwanziger, *supra* note 98, at 2672.

102. *See* Melnick & Zwanziger, *supra* note 98, at 2669.

103. *Id.* at 2670.

104. *Id.* at 2672.

105. *Id.* at 2673.

106. *See* Robinson, *supra* note 98, at 2723. While overall cost reductions were achieved during the period, there was a considerable rate of cost increase per hospital admission during the period. *Id.*

expectation holds true—competition results in lower prices or in lowering the price/cost margin.[107] In a payer-driven market, purchasers are motivated and capable price shoppers.[108] In such circumstances, margins, measured using the bargained-for price rather than the list price, have been shown to have fallen in competitive markets in the period 1983 to 1988 after the introduction of broadened selective contracting.[109] Similarly, California hospitals having more than ten other hospitals within a fifteen mile radius were found to have an adjusted inflation rate of 40.5%, whereas California hospitals having fewer than ten hospitals within a fifteen mile radius had an adjusted inflation rate of 62%. Hospitals in other states had a comparable overall rate of 58.4%.[110] And under selective contracting, the California Blue Cross Preferred Provider Organization (PPO) was able to secure lower prices for its patients in competitive markets.[111]

These empirical studies call into question the claim that antitrust policies are inappropriately applied to the healthcare marketplace because of some purported special characteristics of that market. Since hospitals and other providers apparently can be induced to compete on the basis of a variety of factors, including price, it would seem to be important to maintain potentially competitive markets so that consumers may realize the benefits of price and other forms of competition.[112]

These empirical studies also highlight the importance of evaluating empirical evidence by taking into consideration the structure of the market in which the empirical studies were conducted. These research findings caution against jumping to the conclusion that normal economic expectations are inapplicable in a market when economic theory would predict otherwise. One must be careful to evaluate and apply these empirical findings to situations sensitively, remembering that changes in institutional structure can have significant effects on how participants in the market conduct their affairs once incentives are altered.

Empirical evidence may well be useful in assuring that antitrust enforcement

---

107. *See* Dranove et al., *supra* note 101, at 179, 182; Glenn A. Melnick, *The Effects of Market Structure and Bargaining Position on Hospital Prices*, 11 J. HEALTH ECON. 217, 231-32 (1992); Zwanziger & Melnick, *supra* note 91, at 316; Jack Zwanziger et al., *Cost and Price Competition in California Hospitals, 1980-1990*, HEALTH AFF. Fall 1994, at 118, 123; Zwanziger et al., *supra* note 100, at 429. For more recent data, see Alain C. Enthoven & Sara J. Singer, *Managed Competition in the California Health Care Economy*, HEALTH AFF., Spring 1995, at 105; Ron Winslow, *Is Victory in Sight in Health-Care War?*, WALL ST. J., Feb. 28, 1995, at 1 (attributing a 1.1% drop in average costs per employee from a Foster Higgins survey of employers' shifts to enrollment in managed care plans).

108. *See* Dranove et al., *supra* note 101, at 183; Zwanziger et al., *supra* note 107, at 120; Zwanziger et al., *supra* note 100, at 427-29.

109. *See* Dranove et al., *supra* note 101, at 201.

110. *See* Robinson & Luft, *supra* note 100, at 2679.

111. *See* Melnick, *supra* note 107, at 229, 231.

112. *See* Zwanziger et al., *supra* note 107, at 125; Zwanziger et al., *supra* note 100, at 442-44.

authorities understand what conduct threatens competition and what activities may be procompetitive. This is the Supreme Court's mandate as it has applied antitrust law to various components of the healthcare industry.[113] The doctrine itself seems sufficiently flexible to accommodate these concerns,[114] and the adoption and continued revision and updating of the antitrust enforcement guidelines (safety zones) in this area suggest that antitrust enforcement officials are quite aware of and sensitive to the need to conform enforcement efforts to an evolving understanding of the realities of the medical care marketplace.[115]

C. The Butterworth *Case*

A recent decision of a federal district court denying a preliminary injunction in a hospital merger case brings together the empirical and normative strands in the analysis of antitrust issues in the health care arena. *FTC v. Butterworth Health Corp.*[116] involved the merger of the two largest hospitals in Grand Rapids, Michigan, a four-hospital city.[117] The district court essentially accepted the analysis of the FTC as to the relevant product and geographic markets and concluded that "the proposed merger would result in a significant increase in the concentration of power in two relevant markets, and produce an entity controlling an undue percentage share of each of those markets."[118] Nevertheless, the district court declined to issue the requested preliminary injunction.

In setting forth the appropriate analytical standard, the district court in *Butterworth* stated that the defendant hospitals could defeat the FTC's *prima facie* case by showing that the "proposed merger is not likely to result in anticompetitive effects."[119] While that stated standard retains the proper focus of antitrust analysis—a balancing of procompetitive against anticompetitive effects[120]—it is far from clear that the district court avoided the prohibited pitfall of accepting the argument that the "special characteristics of a particular

---

113. *See* FTC v. Indiana Fed'n of Dentists, 476 U.S. 447 (1 986); National Soc'y of Prof'l Eng'rs v. United States, 435 U.S. 679 (1978); Goldfarb v. Virginia State Bar, 421 U.S. 773, 788-89 n.17 (1975). *See generally* William M. Sage, *Judge Posner's RFP: Antitrust Law and Managed Care*, 16 HEALTH AFF. 44 (1997) (noting need for better empirical evidence in managed care antitrust cases).

114. *See* Meyer & Rule, *supra* note 63, at 182-220.

115. U.S. Department of Justice & Federal Trade Commission, *supra* note 64.

116. 946 F. Supp. 1285 (W.D. Mich. 1996), *aff'd per curiam,* 121 F.3d 708 (Table), text at 1997 WL 420543 (6th Cir. July 8, 1997).

117. *Id.* at 1288.

118. *Id.* at 1294.

119. *Id.*

120. *See, e.g.*, FTC v. Indiana Fed'n of Dentists, 476 U.S. 447, 459 (allowing for consideration of "countervailing pro-competitive virtue[s]").

industry"[121] or organizational form[122] justify "monopolistic arrangements" on the ground that they "will better promote trade and commerce than competition."[123]

The district court purported to use empirical analysis to conclude that the merger should be allowed to proceed. The court seemed influenced by the work of an economist, Dr. William J. Lynk, whose studies concluded that market concentration among nonprofit hospitals was "positively correlated not with higher prices, but with lower prices."[124] Dr. Lynk's research tended to support the "medical arms race" hypothesis, which earlier research on California hospitals had supported.

The FTC criticized the Lynk research on the ground that it did not control for the effects of different levels of costs in concentrated and non-concentrated markets. The FTC argued that labor costs, in particular, were lower in rural areas, where levels of hospital concentration tended to be higher. The differences in observed prices, according to the FTC, could be attributable to differences in costs in concentrated and non-concentrated markets. Since the level of costs facing hospitals was so variable, the FTC argued, the Lynk research did not adequately distinguish among possible contributing elements. A raw correlation study, the FTC contended, was too crude a measure to determine whether it was concentration levels or cost factors that led to the observed pricing results.[125]

The district court noted that these methodological debates would probably continue long after the case was over, but that at the very least there was agreement that "high market concentration among nonprofit hospitals does not correlate positively with higher prices."[126] For the district court, that finding established a "good reason to question the applicability of the traditional presumption that a significant increase in market concentration will lead to higher prices in connection with the merger of nonprofit hospitals."[127]

The reliance of the court in *Butterworth* upon Dr. Lynk's empirical findings was placed within the framework of traditional antitrust analysis. In an earlier merger case, the Seventh Circuit Court of Appeals (per Judge Posner) had noted that antitrust merger cases are "decided on the basis of theoretical guesses as to what particular market-structure characteristics portend for competition," and urged empirical study of the issue.[128] The *Butterworth* court ostensibly viewed

---

121. *See* National Soc'y of Prof'l Eng'rs v. United States, 435 U.S. 679, 689 (1978).

122. *See, e.g.*, American Soc'y of Mechanical Eng'rs, Inc. v. Hydrolevel Corp., 456 U.S. 556, 576 (1982) ("[i]t is beyond debate that nonprofit organizations can be held liable under the antitrust laws.").

123. *See National Soc'y of Prof'l Eng'rs*, 435 U.S. at 689.

124. *Butterworth*, 946 F. Supp. at 1296. *See* William J. Lynk, *Nonprofit Hospital Mergers and the Exercise of Market Power*, 38 J.L. & ECON. 437 (1995).

125. *Butterworth*, 946 F. Supp. at 1295-96.

126. *Id.* at 1295.

127. *Id.*

128. United States v. Rockford Mem'l Corp., 898 F.2d 1278, 1286 (7th Cir. 1990). *See generally* Sage, *supra* note 113.

the Lynk analysis as a response to that suggestion.

To the extent that the *Butterworth* decision rested on empirical findings related to the effect of concentration in a particular market, one could argue that its analysis was in accord with prevailing antitrust doctrine—it was just an attempt to secure a better understanding of how firms in a particular market and with a particular organizational form (nonprofit) would respond to concentrated market conditions. There might be legitimate methodological concerns, such as failing to control for geographical variation in costs,[129] or to consider how the change in market structure seemed to have altered the behavior of California hospitals (many of which were nonprofits) once they faced the reality of selective contracting.[130] Still, the analysis might be taken as within the traditional framework of rule of reason balancing of procompetitive and anticompetitive consequences of certain conduct.

When the two biggest hospitals in a four-hospital market merge, however, that necessarily reduces the number of independent decisionmakers in the market. In the Grand Rapids hospital market, the court acknowledged that there were "substantial barriers to new entry into the relevant market,"[131] thereby posing a significant risk of circumscribed consumer choice. Indeed, the district court found that the two small remaining hospitals' "ability to compete with the merged entity and defeat a small but significant price increase would be limited, especially for the foreseeable future."[132] This finding by the district court makes its purported reliance on empirical evidence within a traditional antitrust analytical framework troubling. It suggests that competition is concededly diminished, but that the institutions who have the capacity to raise price as a matter of market structure would not in fact exercise that power as a matter of noblesse oblige. At that point, the district court's decision subtly but unmistakably transmutes—from a traditional consideration of empirical evidence in a particular market and the structural effect of a merger on competition to a consideration of the much more amorphous question of how specific firms in a market will conduct themselves given their capacity to influence price and output in a market.

The district court's opinion does not, therefore, seem to be based on the procompetitive advantages of the hospital merger under review. Instead, the court implicitly seems to have shifted paradigms—from ascertaining the competitive impact of the proposed merger to determining the totality of overall benefit to the community derived from the merger. The latter inquiry is fundamentally at odds with core Supreme Court teaching on the appropriate analysis in antitrust cases; it impermissibly allows non-competitive values to be weighed against the virtues derived from competition, which is the core value of the antitrust laws.[133] In this sense, then, the court's mode of analysis subtly, but

---

129. *Butterworth*, 946 F. Supp. at 1295.

130. *See supra* notes 82-113 and accompanying text.

131. *Butterworth*, 946 F. Supp. at 1297.

132. *Id.*

133. *See* Blumstein & Sloan, *supra* note 24, at 28-32.

definitively shifts from empirical to normative.

The first element of the court's normative approach is a focus on the nonprofit status of the two merging hospitals.[134] Previous cases had declined to give determinative weight to nonprofit status as a justification for anticompetitive consolidations.[135] For example, in *FTC v. University Health, Inc.*, [136] the Eleventh Circuit noted that, whatever present or past intentions, a nonprofit entity and its governing body would be "free to decide where to set prices and output."[137] Once a nonprofit entity enjoys concentrated authority, its "business decisions are not mandated by law."[138] In *Butterworth*, however, the court seemed to place greater emphasis on the finding that "a substantial increase in market concentration among nonprofit hospitals is not likely to result in price increases."[139] While disclaiming that the merging hospitals' status was dispositive in the matter, the court nevertheless placed considerable emphasis on noblesse oblige considerations associated with the leadership of the nonprofit institutions: "[T]he involvement of prominent community and business leaders on the boards of the hospitals can be expected to bring real accountability to price structuring,"[140] especially in view of the specific enforceable commitments that the hospitals were willing to make for contributions to community betterment.

This approach abandons reliance on the structural guarantees of a competitive marketplace in favor of reliance on alternative mechanisms of assuring accountability. Competition imposes market discipline and is designed to work by establishing a framework of incentives that does not rely on administrative or political mechanisms of enforcement. The antitrust law polices the process of competition, but it is the process of competition itself, accompanied by the economic incentives confronting the participants in the market, that results in a socially appropriate self-policing system.

There are mechanisms through which states can substitute a system of regulation for competition. The *Parker v. Brown*[141] state action doctrine allows

---

134. The authority of the FTC to challenge mergers of nonprofit hospitals had been established in earlier cases. *See, e.g.*, FTC v. Freeman Hosp., 69 F.3d 260, 266-67 (8th Cir. 1995); FTC v. University Health, Inc., 938 F.2d 1206, 1214-15 (11th Cir. 1991); United States v. Rockford Mem'l Corp., 898 F.2d 1278 (7th Cir.), *cert. denied*, 498 U.S. 920 (1990).

135. *See, e.g.*, NCAA v. Board of Regents, 468 U.S. 85, 101 n.23 (1984) ("Good motives will not validate an otherwise anticompetitive practice."); American Soc'y of Mechanical Eng'rs, Inc. v. Hydrolevel Corp., 456 U.S. 556, 576 (1982) ("It is beyond debate that nonprofit organizations can be held liable under the antitrust laws."); Hospital Corp. of Am. v. FTC, 807 F.2d 1381, 1390 (7th Cir. 1986) ("The adoption of the non-profit form does not change human nature."), *cert. denied*, 481 U.S. 1038 (1987).

136. *University Health*, 938 at 1206.

137. *Id.* at 1224.

138. *Id.*

139. *Butterworth*, 946 F. Supp. at 1297.

140. *Id.*

141. 317 U.S. 341 (1943).

states to confer immunity on private conduct that might otherwise run afoul of the federal antitrust law.[142] But *Parker* is a narrowly circumscribed doctrine, which requires that a state clearly articulate a policy to substitute regulation for competition and that the state actively supervise private conduct[143] on an ongoing basis[144] to assure that the policies being pursued by private parties are reflective of state policy and are not just reflective of private norms.[145] Actual, not just potential supervision by accountable state officials is required under *Parker*.[146] In addition, state officials not only must possess ultimate authority to review private decision making, they must exercise it.[147] Passive ratification is insufficient.[148]

The kind of evidence of self-abnegation on the part of the two nonprofit merging hospitals in *Butterworth* is a far cry from the stringent antitrust immunity standards of the *Parker* state-action doctrine. The district court did impose an obligation that the hospitals' commitment to the community to freeze prices and limit margins be embodied in a court decree,[149] presumably providing some enforcement vehicle for those commitments. But the very development of such a plan, which has the hallmarks of the hospital cooperation legislation discussed in Part II.B suggests an accountability framework in considerable tension with or even at odds with the accountability mechanism contemplated by the antitrust law—namely, competition. Further, the substitution of regulatory for competitive mechanisms is permitted to states because of principles of federalism. Yet, no responsible political actor in Michigan clearly articulated a policy to abrogate the norm of competition or actively supervised the private conduct in question.[150]

Perhaps the most troubling element of the district court's analysis is the court's treatment of the managed care issue. This may be the most clear-cut example of how the *Butterworth* analysis is really driven by normative rather than empirical concerns.

Apparently, the FTC contended that a significant impact of the merger in Grand Rapids would be on managed care organizations' ability to win price

---

142. *See* Blumstein, *supra* note 15, at 1486-89.

143. *See* California Retail Liquor Dealers Ass'n v. Midcal Aluminum, Inc., 445 U.S. 97, 105 (1980). The "clear articulation" requirement can be satisfied by a state policy of encouraging or authorizing, though not requiring, anticompetitive conduct. *See* Southern Motor Carriers Rate Conference, Inc., v. United States, 471 U.S. 48, 59-61 (1985).

144. *See* FTC v. Ticor Title Ins. Co., 504 U.S. 621, 633 (1992).

145. *See* Town of Hallie v. City of Eau Claire, 471 U.S. 34, 47 (1985) (With respect to private anticompetitive conduct, "there is a real danger that [the private party] is acting to further [its] own interests, rather than the governmental interests of the State.").

146. *See Ticor Title Ins. Co.*, 504 U.S. at 633.

147. *See* Patrick v. Burget, 486 U.S. 94, 101 (1988).

148. *See Ticor Title Ins. Co.*, 504 U.S. at 633.

149. *Butterworth*, 946 F. Supp. at 1298.

150. Michigan was not one of the states, in fact, that enacted a hospital cooperation law. *See* Blumstein, *supra* note 65, at 253-54.

concessions from the merging hospitals.[151] Greater concentration in the hospital market would allow the merging hospitals "to stem the growing influence of managed care organizations, whose growth has competitively secured discounts from hospitals."[152] The district court acknowledged that the merging hospitals "made no secret" of their desire to standardize managed care rates, resulting in price increases for some managed care organizations and price decreases for others.[153]

Quite understandably, the FTC contended that the hospitals' plan, which the court surrealistically labeled "level[ing] the managed care organization playing field,"[154] would create an anticompetitive result. In a competitive market, each exchange transaction between a willing buyer and a willing seller will reflect the bargaining power of the negotiating parties. It is irrational for a seller to price its product or service at below marginal cost, because such transactions are not profitable, even at large volumes. Thus, one can assume that the Grand Rapids hospitals were not pricing their services to managed care organizations at below marginal cost. However, it is possible that the hospitals were pricing their services to some powerful managed care organizations at below fully allocated average cost. That is, hospitals likely were not pricing their services so low that they were not recovering their variable costs—costs associated with delivering those services. At the same time, it is very possible that some purchasers were making smaller contributions to overhead than were others.

In response to the impetus toward cost containment that stems from increased competition, hospitals traditionally have sought to retain their revenues by shifting costs from some payers onto others.[155] But "[c]ost-shifting strategies can succeed only if (a) other buyers (onto whom costs are shifted) lack market leverage, or (b) transaction costs of prudent, aggressive purchasing by other buyers exceed potential benefits to those buyers in terms of cost savings."[156] This means that cost shifting cannot be a stable long-run condition because it "hinges on a passive payor community."[157] As buyers become more involved in administering their healthcare costs, and as those costs escalate, buyers have the ability and the incentive to manage those costs more effectively. Combining purchasing power through managed care organizations is an example of this phenomenon.

The district court in *Butterworth* was clearly unpersuaded about the virtues of managed care and the effect of tough bargaining by managed care

---

151. *Butterworth*, 946 F. Supp. at 1299.

152. *Id.*

153. *Id.*

154. *Id.*

155. For a discussion of cost shifting, see Charles E. Phelps, *Cross-Subsidies and Charge-Shifting in American Hospitals*, *in* UNCOMPENSATED HOSPITAL CARE: RIGHTS AND RESPONSIBILITIES 108 (Frank A. Sloan et al. eds., 1986).

156. *See* Blumstein, *supra* note 15, at 1480. *See generally* MICHAEL A. MORRISEY, COST SHIFTING IN HEALTH CARE: SEPARATING EVIDENCE FROM RHETORIC 85-89 (1994).

157. Blumstein, *supra* note 15, at 1481.

organizations with the merging hospitals. The court accepted the cost shifting story and assumed that the hospitals could continue to shift costs into the future in a competitive environment. Yet the ability of hospitals to shift costs assumes the existence of unexercised market power over at least a segment of the market and, incidentally, undermines the court's view that nonprofit hospitals do not take advantage of their market leverage. Cost shifting is evidence of the contrary, as hospitals offset projected shortfalls in revenue from competitive purchasers by imposing additional costs on the less powerful, less knowledgeable, or less well-organized purchasers. But this assumption is questionable as a market becomes more competitive and purchasers become more sophisticated.

Because of cost shifting, the district court in *Butterworth* believed that the price discounts obtained by the managed care organizations were "illusory."[158] But those discounts could only be characterized as "illusory" in the sense that the district court is thinking of redistributive values that may be at stake. Relationships among consumers call into play traditional health policy concerns regarding equity and access to medical services of appropriate quality. Those concerns do not reflect considerations regarding efficiency, which is the hallmark or core value of the antitrust laws. The district court, it seems, was worried about the potential loss of cross subsidization from some purchasers to others. The assumption that such cross subsidization could occur over the long term in a competitive market is questionable. It is likely that the hospitals' concern about the stability of that traditional financing scenario led the hospitals to resist the leverage of the managed care purchasers and to seek out ways (such as the merger) to counterbalance the market power of the managed care organizations with market power of their own. Calling the ability of some consumers in a market to secure lower prices "illusory" because of the feared potential impact on other consumers, is nothing short of a challenge to the fundamental antitrust premise that worthy purposes unrelated to competition cannot justify collective conduct that is, on its own, anticompetitive.

In the Grand Rapids market, it was acknowledged that hospital over-capacity existed. Under such circumstances, theory would suggest that some erosion of capital infrastructure would be appropriate. To the extent that a hospital cannot recover its fully allocated average cost, which includes such indirect costs as depreciation, it might be induced to reduce its capacity. This is the market's method of bringing supply into equilibrium with demand. The court's assumption, however, must have been that the hospitals could not downsize as independent actors, but that assumption is unwarranted. Firms faced with overcapacity are frequently faced with the question of how to bring their capacity into synch with the realities of the market. This can be done through reduction in the cost or scope of services delivered and through developing more effective means of delivering services. Alternately, a hospital can seek to preserve or gain market share by developing a specialized niche or providing a superior or less costly service.

---

158. *Butterworth*, 946 F. Supp. at 1299.

The essential point here is that institutions, faced with normal economic incentives, must often cope with overcapacity. The court did not explain satisfactorily why these firms should be able to cope with their overcapacity by exerting greater potential market leverage on purchasers of their service. It is here that the nonprofit status of the merging hospitals seems to have been important.

The district court acknowledged that the two merging hospitals, despite their excess capacity, were operating with above-average profit margins.[159] The FTC contended that the hospitals could cope with their excess capacity by reducing those margins and, implicitly, not seeking to exploit their market leverage by increased cost shifting. The court was unimpressed by this argument because it meant that, with smaller margins, the merging nonprofit hospitals would have less money to reinvest in their facilities. Yet, as the court noted, "[s]uch reinvestment necessarily results in benefits to consumers in the form of expanded and improved services."[160] The court observed that there was no evidence that the hospitals had wasted or otherwise misspent those extra funds.

The district judge himself toured the facilities and concluded that "the boards of both institutions have been responsible stewards of the resources available" and have "continuously reinvested substantial sums in their facilities to keep pace with medical and patient demands."[161] If there were a reduction in resources available to the facilities, that "could only have adverse effects on the quality of care provided."[162] This statement is quite important and warrants more detailed analysis.

Throughout this component of the opinion, the district court expressed concern about the overall quality of care and about consumers who are not members of managed care plans. But these are not the traditional concerns with economic efficiency that underlie antitrust analysis. With respect to consumers who are not members of managed care plans, the court assumed that the hospitals would continue to be able to exercise leverage over them indefinitely. This is far from clear, as the earlier discussion explained. Managed care organizations have arisen with considerable buying leverage; they typically are made up of many employer-groups, and they secure their bargaining power by aggregating demand from all those insured by their plans. There is no reason to believe that, as costs escalate and information about aggregation becomes more readily accessible, the buyers for whom the court expresses concern will not be able to defend themselves in the market. At any rate, it is far from clear that the remedy for that type of overcharging is to allow the overcharging party even more economic leverage.

The cost shifting story quickly gets interwoven with the concern for the fiscal well-being of the merging hospitals. One can reasonably assume that an institution worried about its long-term ability to cost-shift would seek to take

---

159. *Id.*
160. *Id.*
161. *Id.*
162. *Id.*

measures to defend itself against those who have had the market clout to negotiate favorable prices. If cost shifting looms as an evanescent option, then the available alternative strategies for the hospital are all somewhat unattractive. These include reducing the level of cross-subsidized services, finding less costly ways of providing services through efficiencies (such as changing the skill mix among employees, reducing salaries of some of the higher paid employees, or reducing staff levels overall), or lowering the quality of services to approximate what the buyers are willing to pay for. The antitrust law has rejected a "worthy purpose" defense against anticompetitve conduct to achieve professionally desirable outcomes.[163] It is highly questionable whether the antitrust law should allow the compromising of effective competition for well-situated consumers to achieve such a worthy purpose as high quality medical care or to achieve other worthwhile goals pursued by nonprofit hospitals. In a market, levels of quality are driven by the purchasers, not by the professionals in charge of the hospital. This is an essential ingredient of the market system: purchasers control price, quality, and output. In medical care, however, under the professional model, the tradition has been otherwise: health care suppliers control price, quality, and output.

That is where the antitrust law's normative importance enters the analysis. The focus on the values of efficiency and purchaser control assures the primacy of market-oriented values. The district court opinion reflects an adherence to the professional paradigm and a willingness to bend the antitrust law to soften the impact of the economic marketplace with consideration of professional values such as quality of care without regard to cost and with consideration of traditional redistributive values of medical care concerning the plight of less organized consumers.[164]

A nice contrast with the revanchist approach of the district court in *Butterworth,*[165] is the district court's decision in *United States v. Long Island Jewish Medical Center (LIJ).*[166] Whereas *Butterworth* seemed driven more by normative than empirical considerations and seemed to embrace a paradigm for the structure of the healthcare industry that is in tension with the enforcement of the antitrust laws, the district court in *LIJ* appeared to use empirical evidence in a more traditional antitrust analytical framework.

LIJ involved a proposed merger of two nonprofit teaching hospitals on Long Island, New York. In denying the FTC's request for a preliminary injunction to block the merger, the district court examined the empirical evidence regarding traditional criteria such as the definition of the product and geographic markets. The district court believed that the FTC had not established its market

---

163. *See* Blumstein & Sloan, *supra* note 24, at 28-32.

164. For a comprehensive discussion of the different ways of thinking about medical care and the bulwarks and inroads on the professional model, see Blumstein, *supra* note 15, at 1463-86.

165. The truncated per curiam affirmance of *Butterworth* by the court of appeals, unfortunately, did not give the case the in-depth consideration it warranted. No. 96-2440, 1997 WL 420543 (6th Cir. July 8, 1997).

166. No. CV97-3412 (ADS), 1997 WL 662731 (E.D.N.Y. Oct. 23, 1997).

definitions given the expert and empirical information presented and concluded that there would not be an unwarranted lessening of competition from the merger.[167]

Although a number of the same arguments on behalf of the merging hospitals were made in *LIJ* and in *Butterworth*, the court's treatment of those arguments was much more in accord with the traditional antitrust framework. The district court's handling of *LIJ* is a better reflection of how empirical evidence should inform antitrust analysis, as the court appropriately took into account data regarding cross elasticity of demand and evidence regarding the scope of product and geographic markets. With regard to the difficult issue of nonprofit status, the court evaluated the evidence as part of its overall analysis of the competitive impact of the merger but clearly gave "only limited and non-determinative effect" to that factor.[168] And importantly, the court rejected on empirical grounds the FTC's claim that entry into the market was circumscribed.[169] The court found that, under appropriate government guidelines, one hospital fit the criteria as a likely entrant that would impose market discipline on the merging hospitals.[170]

The differences in analytical approach between *LIJ* and *Butterworth* led to the same result; in both cases, the courts denied the preliminary injunction that the FTC sought. Where the *LIJ* court used empirical analysis within a traditional antitrust framework, the *Butterworth* court seemed to push the envelope of antitrust enforcement with an adherence to a paradigm of the healthcare industry that is, at least, in tension with the pro-market mandate of antitrust law and, at most, fundamentally inconsistent with the dictates of antitrust law. Therefore, where *LIJ* may reflect a substantive doctrinal review of the tools and approaches for enforcement of the antitrust law, it does not threaten the procompetitive normative premises of the antitrust law itself. Unfortunately, the mixed messages that emanate from *Butterworth* may undermine the ability of the enforcement agencies to apply the procompetitive policies of the antitrust law—for all their substantive and symbolic importance—to an important component of the healthcare marketplace.

---

167. *Id.* at *26.
168. *Id.* at *28.
169. *Id.* at *31.
170. *Id.*

RECOGNITION OF HEALTH CARE MARKET ANOMALIES

Comments on a Paper by Professor James Blumstein

JOHN C. RENDER*

Professor Blumstein has presented a provocative and thorough overview of the rapidly changing paradigms in the delivery of health care in the United States and the effect that empirical evidence is having on these market-driven forces, the study of which is now much in vogue.[1]

I have little doubt that the free market approach to the allocation of resources in the health care field,[2] where structured properly, is probably the best method to achieve effective cost competition and quality products. Further, contrary to the view expressed by some authors,[3] I believe that the health care field reacts and responds to many, but not all, traditional market forces endemic to the free market system.

For example, providing financial incentives or disincentives clearly affects the provision of services, the utilization of services, the methods used to provide services, and probably the quality of the product.[4] The enactment of the Medicare prospective payment system for hospitals in 1983[5] and the expansion of Medicare coverage for such services as renal dialysis, home health care and others are merely the most noteworthy illustrations that the behavior of health care providers, and consumers, will radically change if money is withheld, as in the former example, or is provided as in the latter instances.

However, these examples serve only as indications that some market forces do function in the health care field in a manner similar to other markets where

---

* Partner, Hall, Render, Killian, Heath & Lyman. P.C. B.S., 1966, Butler University; J.D., 1971, Indiana University School of Law—Indianapolis.

1. James F. Blumstein, *The Application of Antitrust Doctrine to the Healthcare Industry: The Interweaving of Empirical and Normative Issues*, 31 IND. L. REV. 91 (1998).

2. WEBSTER'S NINTH NEW COLLEGIATE DICTIONARY 491 (1987) defines free market as "an economic market operating by free competition." This necessarily implies an absence of government regulation or intervention. *See generally* Andrew Farris & Griffin Seiler, *Health Care Reform: A Free-Market Proposal*, 7 LOY. CONSUMER L. REP. 45, 45-46 (1995) (positing that "[t]he fundamental problems in the health care market are a result of government intervention" and proposing free-market reform which vests health care choice and responsibility directly in the consumer).

3. *See, e.g.*, Bengt Jönsson, *What Can Americans Learn from Europeans?*, 2 HEALTH CARE FINANCING REV. 79, 91-92 (Supp. 1989); Paul Starr, *The Framework of Health Care Reform*, 329 NEW ENG. J. MED. 1666, 1668 (1993).

4. *See* Robert H. Brook et al., *Does Free Care Improve Adults' Health? Results from a Randomized Controlled Trial*, 309 NEW. ENG. J. MED. 1426, 1432 (1983); Alan L. Hillman, *Financial Incentives for Physicians in HMOs: Is There a Conflict of Interest?*, 317 NEW ENG. J. MED. 1743, 1748 (1987).

5. Social Security Amendments of 1983, Pub. L. No. 98-21, § 601, 97 Stat. 65, 149-72 (codified as amended at 42 U.S.C. § 1395ww (1994 & Supp. I 1995)).

there is a desired product and consumers available to use the product and pay for it (or as is often the case in health care, have someone else pay for it).[6]

Professor Blumstein strongly argued in his presentation that it is far too early in the life of the market-based paradigm in health care to consider radical surgery and that efforts to craft special exceptions to a market system because of the uniqueness of health care were largely based on anecdotal evidence or outright hostility to market driven initiatives.[7] He cited the fairly widespread enactment of state hospital cooperation laws as an indication that incomplete data and strongly held beliefs, even if wrong, are a potent combination which should be resisted until there is better data and more significant education of the public as to the value of market driven systems.[8]

One of the major barriers to pro-competitive conduct among health care professionals has been a particularly long history of cooperation and collaboration encouraged by perceived community benefit.[9] Further, these joint activities of competitors may have even been lawful prior to *Goldfarb v. Virginia State Bar*.[10] Since *Goldfarb*, it has been difficult for many providers of health care to change their cooperative ways and to view competitive initiatives as being public minded and in the best interests of their communities.

In fact, purchasers of health care, legislative bodies, community leadership and other important constituencies of health care have in many and varying ways urged a continuation of the traditional policies of cooperation and collaboration.[11] Purchasers, for example, when discussing this topic will suggest that they desire

---

6. Recent surveys indicate that over 80% of Americans have health insurance. Karen Donelan et al., *Whatever Happened to the Health Insurance Crisis in the United States? Voices from a National Survey*, 276 JAMA 1346, 1347 (1996). Nevertheless, 16% of the population said they had a problem in the past year paying medical bills, with the uninsured being three times more likely than the insured to give that response. *Id.*

7. Blumstein, *supra* note 1, at 91.

8. *Id.* at 102.

9. *See* Clark C. Havighurst, *Regulation of Health Facilities and Services by "Certificate of Need,"* 59 VA. L. REV. 1143, 1148-51 (1973).

10. 421 U.S. 773 (1975). *See generally* John D. Blum, *A Consumer Perspective on the Pros and Cons of Antitrust Enforcement in Health Care: An Introduction*, 8 LOY. CONSUMER L. REP. 76 (1995-1996) (noting that *Goldfarb* "opened the doors for the application of the antitrust laws in the health care field.").

11. *See generally* James F. Blumstein, *Assessing Hospital Cooperation Laws*, 8 LOY. CONSUMER L. REP. 98, 103 (1996) (noting that at least 19 states allow cooperative agreements among hospitals by immunizing behavior that otherwise might be subject to federal antitrust scrutiny); James M. Lasley, *Hold Hospital Directors Accountable*, FLA. TODAY, Sept. 3, 1996, at 6A (letter to the editor in which the author writes "the duplication of very expensive medical services; i.e., building a new hospital within miles of another hospital, is completely beyond absurd."); Stephen P. Bunker, *Community Health Care Is a Top Priority*, FLA. TODAY, Sept. 10, 1996, at 6A (letter to the editor in which hospital administrator agrees with the previous letter that "[d]uplication of services does nothing to improve the health [of the community]. . . . In fact, it increases costs to the community.").

providers in a given service area to collaborate and not duplicate unnecessary services and equipment, but only to the extent that meaningful choices for consumers and bargaining alternatives for purchasers remain available. In other words, collaboration and cooperation should continue only to a point of mergers and consolidations which limit bargaining ability on the part of purchasers and choice on the part of consumers. This is a difficult concept for many managers in health care to embrace and still avoid the innumerable anti-trust land mines inherent in such collective endeavors.

A characteristic of the health care field that distinguishes it from most other commodities and services is the provision of a vital human service. Most other vital human services such as water and power are subject to state oversight in the form of public utility commissions or similar entities. This is based, at least in part, on the notion that such services are so significant that determining their availability by market forces is contrary to civilized values and should therefore not be subject to the varieties of the market system. While it is not clear that health care fits neatly into a box wherein a regulatory scheme is the only way in which resources can be allocated, it certainly has some characteristics which would suggest that solution.[12]

In considering the proper role of "the market" in the delivery of health care, empirical evidence can play an important role in framing public policy. Several important issues, however, should be considered.

(1) Health care is an area in which the public has considerable experience and perceived knowledge. This may make it more difficult to make policy decisions based largely on empirical data since the greater the knowledge or experience of the public in a policy area, the more arduous it is to formulate public policy based upon dispassionate empirical data. For example, duplication of services and equipment by health care providers is not always seen by the consuming public as being pro-competitive nor positive.[13] Such duplication of competing services may not result in better quality or lower prices.[14] Thus, there is considerable public support for the continuation of the existing cooperation and collaboration among health care providers in their community, particularly for tertiary and highly technical and expensive services. Thus, while empirical data might show competition is the best allocator of health care services, personal experience and convictions of the public may still greatly influence ultimate policy.

(2) Health care markets can be very imprecise economic models and often

---

12. Consumers of health care oftentimes possess limited information, and their day-to-day choices are similarly limited by the dictates of their health insurance policy. Because the vast majority cannot finance their health costs out-of-pocket, however, they must rely on health insurance. Eliot Freidson, *The Centrality of Professionalism to Health Care*, 30 JURIMETRICS J. 431, 438 (1990).

13. *E.g.*, Lasley, *supra* note 11.

14. *E.g.*, Bunker, *supra* note 11. Significant efficiencies can result from hospital collaboration. *See generally* David Dranove et al., *Is Hospital Competition Wasteful?*, 23 RAND. J. ECON. 247 (1992).

very non-traditional, thus limiting the success of market-driven solutions.[15] For example, the public demand for health care goes beyond desire or need and is widely considered to be a right, not merely a privilege.[16] Further, federal and state laws require providing health care services irrespective of the user's ability to pay for such services.[17]

(3) Unlike many consumer products, health care is not readily capable of qualitative measurement. This lack of comparability to widely accepted standards of quality and value is particulary noteworthy as the service modality increases in complexity. Limited tools exist to determine the quality of the diagnostic skills of completing endocrinologists, for example.

(4) The isolation of the consumer from the economic consequences of purchasing health care in most instances is a clear departure from most economic models in the free market.[18] Apart from deductibles and co-pay provisions in many third-party payment plans, there is often little to deter the consumer from obtaining the product even beyond necessary usage. Contrast this with the provision of food to individuals of need. Where the government provides this assistance, the consumer is given food stamps and remains a direct purchaser accountable for the prudent use of limited resources.

(5) The increasing inability of providers of health care services to charge consumers for the proportionate total economic cost of providing care to the providers' universe of patients presents significant problems. The idea of "cost shifting" from one patient class to another is becoming a thing of the past.[19] Often purchasing groups want to pay only "their share" for services provided.

---

15. *See* Terese Hudson, *Mirror, Mirror*, HOSPITALS & HEALTH NETWORKS, April 5, 1996, at 24, 26.

16. *See* James F. Blumstein, *Health Care Reform: The Policy Context*, 29 WAKE FOREST L. REV. 15, 33 (1994).

17. The Hospital Survey & Construction (Hill-Burton) Act, Pub. L. No. 79-725, 60 Stat. 1040 (1946) (codified as amended at 42 U.S.C. §§ 291a to 291o-1 (1994)), provided federal financing for the construction of health care facilities while also requiring the provision of necessary medical services to those unable to pay. *See* BARRY R. FURROW ET AL., HEALTH LAW: CASES, MATERIALS AND PROBLEMS 628 (2d ed. 1991).

Several state statutes also require the provision of medical care to all person, regardless of their ability to pay. *See, e.g.*, ARIZ. REV. STAT. ANN. § 36-2905 (West 1993) (defining "medically needy" persons and setting forth procedures by which they can obtain state-assisted care); CAL. WELF. & INST. CODE § 16704.1 (West 1991) (stating that "[n]o fee or charge shall be required of any person before a county renders medically necessary services . . . ."); S.D. CODIFIED LAWS ANN. § 28-13-32.3 (Michie Supp. 1997) (noting that a "medically indigent" person can receive assistance for the cost of hospitalization by applying to his or her county of residence).

Both the Emergency Medical Treatment and Active Labor Act, 42 U.S.C. § 1395dd (1994), and provisions within provider agreements, 42 C.F.R. § 489.24 (1995), require hospitals to provide emergency care irrespective of a patient's inability to pay.

18. *See supra* notes 6 and 12.

19. *See* Jack Hadley, *Financial Pressure and Competition: Changes in Hospital Efficiency and Cost-Shifting Behavior*, 276 JAMA 1010 (1996).

Thus, few classes of purchasers are willing to provide the resources necessary to care for those persons who are uninsured or otherwise unable to pay for needed health care services.

Ultimately to determine whether the market paradigm in medical care should be fully embraced, it must be asked whether the market should be the sole determinant as to the provision of this basic human commodity. There are some who would have considerable discomfiture with a complete surrender to the market to determine who receives life saving treatments and who does not.[20]

A possible solution is to permit the market to allocate health care resources in areas where reasonable competitive models exist or can exist and to permit some accommodation to the market in those areas where competition probably cannot occur. It is in this latter environment where empirical evidence may be most useful in providing education to policy makers as to when and where such alternatives to the market should be developed. Changing the existing health care system to encourage market initiatives while addressing systemic structural deficiencies will require considerable information being provided to the public since they will be instrumental in influencing legislative bodies. Health care policy to date has often been based on anecdotal evidence, personal experiences, and often the relative influence of various interest groups.[21]

Because of the rapid and complex changes occurring in the delivery of health care in the United States, neither a rigid application of the usual anti-trust principles nor the creation of special immunities or exemptions should be the standard. It seems premature in the life cycle of the market-based health care system to rush to judgment, whether that judgment is in favor of a complete unfettered market-driven system or a regulatory scheme based on comprehensive state or federal oversight.

Empirical data has not yet been a significant tool in driving public policy in those areas where legislators have considerable experience or deeply held convictions. However, there is some promise that particularly ill conceived public policy may be modified by public education based on valid empirical data. Thus, empirical evidence might still serve a major role in crafting necessary refinements in the current market-based paradigm.

---

20. *See supra* note 3 (neither author supports a complete surrender of health care to the free market).

21. *See* James F. Blumstein, *Distinguishing Government's Responsibility in Rationing Public and Private Medical Resources*, 60 TEX. L. REV. 899, 911 (1982).

PRESUMPTIONS, DAMN PRESUMPTIONS AND ECONOMIC THEORY: THE ROLE OF EMPIRICAL EVIDENCE IN HOSPITAL MERGER ANALYSIS

MICHAEL S. JACOBS*

INTRODUCTION

It is hardly coincidental that the quote in the title of this symposium ("lies, damned lies and statistics") should have been applied disparagingly by a federal district court to the ambiguous role of economic evidence in antitrust litigation.[1] Over the past twenty years, economics has provided the exclusive theoretical basis for antitrust enforcement guidelines,[2] preoccupied antitrust scholarship, and served as the predominant tool for judicial analysis of antitrust disputes. Despite the preeminent role of economic theory in antitrust law, empirical economic evidence has played a relatively small part in both the formation and refinement of economic theory and the judicial resolution of individual disputes.

The reasons for the subordinate role of empirical economic evidence in antitrust analysis are not all that clear. Some may be attributed to the power of certain forms of economic theory, whose persuasiveness rests on assumptions no longer, if ever, deemed worthy of careful examination. Others seem to follow from the limitations of the judicial process and its consequent inability to provide suitable conditions for empirical examination. Still others may stem from the difficulty of measuring the relevant portion of rapidly changing markets with the kind of accuracy and timeliness thought necessary to litigation. Whatever the precise constellation of causes, it is well recognized, even by leading adherents of the law and economics approach to antitrust analysis, that the relationship between economic theory and economic "fact" is tenuous at best.[3]

This is not necessarily suggesting that empirical data can or should play a larger role in antitrust analysis than it already does. Except at the largest levels of generality, levels too large either to help refine theory or to prove useful in litigation, the economic theories underlying antitrust law appear oddly oblivious

---

* Professor of Law, DePaul University College of Law; B.A., 1968, Dartmouth College; J.D., 1971, Yale Law School; M.P.H., 1987, Johns Hopkins University.

1. *See* McGahee v. Northern Propane Gas Co., 858 F.2d 1487, 1496 n.19 (11th Cir. 1988).

2. See, for example, the various incarnations, since 1982, of Merger Guidelines and the recent proposal regarding the use of efficiencies evidence in merger analysis. (The FTC's recent recommendations for modifying antitrust policies included the proposal that efficiencies evidence relevant to the issue of a merger's probable competitive effects be admissible in court.).

3. *See, e.g.*, Frank H. Easterbrook, *Allocating Antitrust Decisionmaking Tasks*, 76 GEO. L.J. 305, 308-09 (1987) (declaring that "[t]he empirical foundation on which much antitrust policy was built has been washed away"); *see also* Fishman v. Estate of Wirtz, 807 F.2d 520, 563 (7th Cir. 1986) (Easterbrook, J., dissenting) (discussing merger law in particular and stating that "[p]ropositions about the economics of mergers often are filled with ifs and maybes; competing schools of thought produce different prescriptions"). *See generally* Michael S. Jacobs, *An Essay on the Normative Foundations of Antitrust Economics*, 74 N.C. L. REV. 219, 250-54 (1995).

to data. Many examples of this curious detachment exist. Thus, while observations and studies of single-firm behavior may demonstrate that for-profit firms generally seek to maximize profits (or at least some for-profit firms; economists continue to debate this point),[4] the evidence about profit-maximizing behavior, or absence thereof, by not-for-profit firms is at best inconclusive.[5]

By the same token, merger law generally proceeds on the assumption—an assumption rooted squarely in economic theory—that antitrust should prevent markets from becoming unduly concentrated. Theory tells us that in such markets, one of two anticompetitive scenarios is likely to play itself out: in the merger-to-monopoly scenario, consumers will face a firm with the power to unilaterally raise price profitably above competitive levels; while in the merger-to-oligopoly scenario, the reduced number of firms remaining and in the post-merger market will find it easier to collude successfully (that is, without being detected) to fix prices at supra-competitive levels. However plausible these assumptions may be on an intuitive basis, their empirical foundation is hard to trace and bereft of recent corroboration.

Though antitrust law contains many theories and assumptions with weak or non-existent empirical support, the two described above have not been selected at random. This paper concerns itself generally with hospital mergers and, in particular, with the case of *FTC v. Butterworth Health Corp.*,[6] a recent controversial hospital merger decision characterized by a respected scholar as "the most revolutionary hospital merger decision yet issued . . . ."[7] This is a decision that owes its notoriety in large measure to the novel manner in which it approached various types of empirical evidence offered in defense of the Butterworth merger but not normally found (and if found, not normally accepted) in hospital merger cases. In discussing *Butterworth*, this paper focuses attention on the poorly understood relationship among antitrust merger theory, its application to particular hospital mergers, and empirical data about the business

---

4. *See, e.g.*, RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 17 (4th ed. 1992). *But cf.* Oliver Hart, *An Economist's Perspective on the Theory of the Firm*, 89 COLUM. L. REV. 1757, 1758-59 (1989) (stating the view that profit maximization may become secondary to the goals of corporate managers).

5. *See, e.g.*, Herbert Hovenkamp, *Positivism in Law & Economics*, 78 CAL. L. REV. 815, 827-31 (1990) (asserting that although the profit-maximizing assumption lacks empirical validation and is difficult to test empirically, it is a normative convention adopted by economists without empirical proof). Others have theorized that the limits of human knowledge and a natural inclination to follow proven patterns of behavior incline (even for-profit) firms to act less than optimally, "satisficing" instead of maximizing. *See* ADOLPH A. BERLE & GARDINER C. MEANS, THE MODERN CORPORATION AND PRIVATE PROPERTY 113-16, 303-08 (1967); Robert G. Harris & Thomas M. Jorde, *Antitrust Market Definition: An Integrated Approach*, 72 CAL. L. REV. 1, 28-29 (1984); Herbert Simon, *Rational Decision Making in Business Organizations*, 69 AM. ECON. REV. 493, 503 (1979).

6. 946 F. Supp. 1285 (W.D. Mich. 1996), *aff'd per curiam*, 121 F.3d 708 (6th Cir. 1997).

7. Thomas L. Greaney, *Night Landings on an Aircraft Carrier: Hospital Mergers and Antitrust Law*, 23 AM. J.L. & MED. 191, 212 (1997).

behavior of hospitals generally.

I. THE HOSPITAL MERGER MOVEMENT OF THE 1990S

It is no coincidence that a hospital merger case should provide the focus of this paper. Hospital mergers are perhaps the most dramatic and unsettling evidence of the organizational upheavals that have characterized healthcare finance and delivery over the past five years.[8] Over the last three years alone, the healthcare field has witnessed an unprecedented number of mergers involving hospitals. According to statistics compiled by the American Hospital Association (the primary organization tracking hospital mergers) there were thirteen hospital mergers in 1990, twenty-three in 1991, fifteen in 1992 and eighteen in 1993.[9] In 1994 there was a quantum leap in activity, as more than 650 hospitals were involved in mergers or acquisitions;[10] and, in 1995, that number rose to 735.[11]

The reasons behind this unprecedented merger movement are manifold. Their common denominator, however, is the desire to lower the costs of hospital operation and thus the prices of hospital services. Seen from today's cost-conscious perspective, hospitals suffer from three significant problems. The first is the existence of tremendous excess capacity caused in large part by the governmentally subsidized construction and expansion of the 1950s and 60s. Among other things, these subsidies resulted in more hospital beds, equipment and services that now seem necessary for the size and health of our population. Recent data, for example, place the national occupancy rate for hospital beds at around fifty percent.[12] The second problem is the high cost of hospital operation. Hospital operation costs consume the lion's share of the health insurance dollar, accounting for approximately forty percent of personal health care spending.[13] The third problem is the growing number of low-cost alternatives to high-priced hospital care. Free-standing outpatient facilities offer many procedures—from complex eye surgery, to abortions, to emergency care—that used to be the sole province of hospitals. Doctors' offices provide more complicated services than ever before and mini-hospitals offer birthing centers and a wide range of other

---

8. *See, e.g.*, Robin E. Remis, *Health Care and the Federal Antitrust Laws: The Likelihood of a Harmonious Coexistence*, 13 J. CONTEMP. HEALTH L. & POL'Y 113, 113-14 (1996).

9. Sandy Lutz, *Let's Make a Deal*, MOD. HEALTHCARE, Dec. 19, 1994, at 47.

10. *Id.*

11. Sandy Lutz, *Mergers and Acquisition Report; 1995: A Record Year for Hospital Deals*, MOD. HEALTHCARE, Dec. 18, 1995, at 43.

12. *See* Peter M. Sullivan, *N.Y.'s Medicaid Cuts are Irresponsibly Deep*, NEWSDAY, Feb. 8, 1996, at A47; Ugur Yavas & Donald J. Shemwell, *Competing for Patients and Profit; Analytical Framework Can Help, Marketers Determine the Competitive Strengths and Weaknesses of Hospitals*, J. HEALTH CARE MARKETING, June 1, 1996, at 30.

13. United States v. Mercy Health Servs., 902 F. Supp. 968, 981 (N.D. Iowa 1995), *vacated*, 107 F.3d 632 (8th Cir. 1997).

specialized services.[14] These competing facilities are equipped with the same sophisticated technology available in hospitals, but because they are small and specialized, and usually choose not to assume responsibility for any community-wide, uncompensated care, their costs and prices are lower than those of their hospital counterparts.[15]

With the continued growth in the membership of managed care organizations and the heightened attention of large, self-insured employers to health insurance costs, these problems have made hospitals prime targets for hard bargaining. In a concerted effort to lower the price of health insurance premiums, managed care has reduced hospitalizations generally, reduced lengths of stay, and turned whenever possible to lower cost outpatient options. Hospitals wishing to receive managed care dollars have had to compete not only with this new group of lower cost rivals but also with each other. Managed care has demanded lower prices from hospitals, pitting neighboring hospitals against one another in low-priced bidding wars and, given managed cares' willingness to have patients travel farther for less expensive care, against other hospitals once considered too remote to be competitive. To survive in this climate hospitals have had to find ways to reduce their historically high costs.

Merger is a time-honored method of cost reduction. By consolidating two separate firms into one larger organization, merger permits a variety of cost savings. Increased size will afford the new firm purchasing and borrowing economies. Its labor force can be reduced. Two separate administrations, law firms, accountants, and public relations firms will no longer be necessary. Expensive equipment need not be purchased twice, and capital expenditures—not only for equipment but for structures as well—can thus be lowered significantly. Moreover, when hospitals consolidate, they no longer need to compete with each other on every imaginable front: each can concentrate on those services and procedures that it does best—most successfully and at the lowest cost—raising their patient volumes in these areas and thus lowering their costs and improving their outcomes. These and other potential benefits normally flow from mergers.

Merger is also, however, a time-honored method for accumulating market power as a prelude to raising prices. In an important sense, the federal antitrust laws were enacted in 1890 and first amended in 1914 in response to public concern about the series of mergers and acquisitions near the end of the nineteenth century that created Standard Oil, many of the large railroad systems, and other trusts that were thought to prey on consumers and rivals alike.[16] Hospital mergers have aroused similar concerns in some quarters, touching off fears that the recent wave of consolidation will result in higher prices for consumers and higher profits for hospitals themselves. These concerns have

---

14. *See* Joe Sims, *A New Approach to the Analysis of Hospital Mergers*, 64 ANTITRUST L.J. 633, 639 (1996).

15. *See* Michael S. Jacobs, *When Antitrust Fails*, 71 WASH. L. REV. 899, 902 (1996).

16. *See, e.g.*, United States v. Reading Co., 226 U.S. 324 (1912), *modified by* 228 U.S. 158 (1913); United States v. Standard Oil Co., 173 F. 177 (E.D. Mo. 1909), *aff'd*, 221 U.S. 1 (1911). *See also* Richard Posner, *100 Years of Antitrust*, WALL ST. J., June 29, 1990, at A12.

prompted the antitrust enforcement agencies of the federal government, the Federal Trade Commission (FTC) and the Antitrust Division of the Department of Justice (DOJ), to challenge several recent mergers (including Butterworth) applying to them the same untested presumptions that underlie merger enforcement generally.

II. THE RUDIMENTS OF MERGER LAW AND THEIR UNEASY FIT WITH HOSPITAL MARKETS

A. Merger Analysis

Section 7 of the Clayton Act[17] provides the primary statutory basis for merger enforcement. As enacted in 1914 and amended in 1950, it proscribes mergers whose effect "in any line of commerce or . . . in any section of the country . . . may be substantially to lessen competition, or to tend to create a monopoly."[18] The language of the statute—particularly the phrases "may be"and "tend to"—make its application predictive in nature, directed at the likely future instead of the presumably better known past. Indeed, unlike other antitrust laws whose legislative history allows no confident conclusions about their intent, the merger laws were designed to arrest anticompetitive concentration "in its incipiency," before its effects might be fully, or even partially, known.[19]

The fear of concentration reflected in the Clayton Act arguably has some empirical underpinning, but it is of a casual, political type rather than of a scientific nature. In 1950, when the Act was amended to increase its scope and efficacy, Congress was all too aware of the close relationship between big business and big government that had characterized and helped to produce fascist Germany and Japan. Strengthening federal merger law in order to stifle large increases in aggregate concentration was seen as a means of preserving and fostering ideals of democracy and fair play.[20] With the passage of time, of course, this particular set of post-war concerns has mostly, if not entirely, abated. Nevertheless, the Clayton Act remains unchanged, in part because its underlying political message about the dangers of corporate size and power continue to strike responsive chords with many, and in part because its implicit economic theory about the relationship between concentrated markets and high prices, though essentially untested, seems plausible. Though evidence in support of this theory is inconclusive—increased concentration has brought higher prices for air travel but lower ones for cola and computer chips—the Clayton Act operates on the assumption that economic evils such as higher prices, reduced output, and

---

17. 15 U.S.C. § 18 (1994).

18. *Id.*

19. *See* Brown Shoe Co. v. United States, 370 U.S. 294, 315-23 (1962); *see also* Alan A. Fisher & Robert H. Lande, *Efficiency Considerations in Merger Enforcement*, 71 CAL. L. REV. 1580, 1588-93 (1983) (discussing the legislative history of Section 7).

20. *See, e.g.*, Robert Pitofsky, *The Political Content of Antitrust*, 127 U. PA. L. REV. 1051 (1979).

inferior quality lurk in concentrated markets. Courts have historically applied the Act on the basis of this assumption.

In 1963, in *United States v. Philadelphia National Bank*,[21] the United States Supreme Court announced the basic methodology for the judicial analysis of horizontal mergers and acquisitions, those between direct competitors in the same product or service and geographic markets.[22] Under this approach, which is still favored today, courts must first define the relevant product and geographic markets in which the merging firms compete.[23] Since the purpose of the law is to prevent undue concentration in any market ("line of commerce"), one must first know the market before the merger's effect on concentration can be gauged. Having defined the relevant market, courts are to presume the illegality of any merger producing a "significant increase in the concentration of firms" within the market and resulting in a firm with an "undue percentage share" of the market.[24] As originally conceived, the rule of presumptive illegality applicable to mergers resulting in undue concentration was a very powerful one, rebuttable only by evidence clearly demonstrating "that the merger is not likely to have such anticompetitive effects."[25]

Over time, the power of this presumption has weakened, as courts and regulators have sought to accommodate antitrust theory to the complex reality of a large and diverse economy. Thus, in 1974, the Supreme Court announced in another merger case that a showing of undue concentration could be rebutted by evidence that the government's data regarding market share failed adequately to reflect the defendants' true competitive position.[26] More recently, courts have increasingly come to discount market share as the exclusive determinant of market power when the defendant can prove that structural factors peculiar to the market effectively diminish the power that large market shares theoretically confer.[27] At the same time, the Federal Horizontal Merger

---

21. 374 U.S. 321 (1963).

22. *Id.* at 356.

23. *Id.* at 356, 363; *see also* Philip Areeda, *Market Definition and Horizontal Restraints*, 52 ANTITRUST L.J. 553 (1983); Robert G. Harris & Thomas M. Jorde, *Market Definition in the Merger Guidelines: Implications for Antitrust Enforcement*, 71 CAL. L. REV. 464 (1983); William M. Landes & Richard A. Posner, *Market Power in Antitrust Cases*, 94 HARV. L. REV. 937 (1981).

24. *Philadelphia Nat'l Bank*, 374 U.S. at 363-64 & n.40 (competition is substantially lessened when the combined share of the merging firms would exceed 30% of the relevant market and the post-merger concentration level of the five largest firms in that market was approximately 78%).

25. *Id.* at 363.

26. United States v. General Dynamics Corp., 415 U.S. 486, 501 (1974).

27. *See, e.g.*, United States v. Baker Hughes, Inc., 908 F.2d 981, 985 (D.C. Cir. 1990) ("[T]hat a variety of factors . . . can rebut a prima facie [merger] case has become hornbook law." Among those factors are ease of entry into the post-merger market, the prospect of efficiencies from the merger, excess capacity, degree of product homogeneity, marketing and sales methods, industry structure, weakness of data underlying the prima facie case, high elasticity of industry demand, and high cross-elasticity of supply and demand.) (citing PHILLIP E. AREEDA & HERBERT HOVENKAMP,

Guidelines,[28] which describe the methods by which enforcement agencies analyze mergers and select cases for prosecution, have been loosened to allow for agency consideration—even with mergers in "highly concentrated markets"—of non-statistical factors that might overcome the presumption flowing from undue concentration.[29] Indeed, so many factors have recently been added to the market share presumption that some courts view it as practically devoid of explanatory power,[30] while commentators complain that merger analysis has lost its theoretical underpinning and now regards all data as arguably relevant.[31]

Much more could be said about the growing complexity and increasing confusion that have come to characterize merger analysis generally; but a fuller treatment of that topic is beyond the scope of this paper. It is sufficient to make two final observations First, merger law continues to grow ever more complicated, as new ideas persuade enforcement agencies to expand their analyses into relatively uncharted territory. In the past few years, for example, the FTC has decided that otherwise unobjectionable mergers might nevertheless stifle competition in "innovation markets," markets for the research and development of new products and processes.[32] Though the scope of these markets appears difficult to define and the notion of innovation markets seems inherently indistinct, the FTC has employed this new concept in analyzing the

---

ANTITRUST LAW ¶¶ 919, 920.1, 921, 925, 934-35, 939, at 813-23 (Supp. 1989); HERBERT HOVENKAMP, ECONOMICS AND FEDERAL ANTITRUST LAW § 11.6, at 307-11 (1985); LAWRENCE A. SULLIVAN, HANDBOOK OF THE LAW OF ANTITRUST § 204, at 622-25 (1977)). *See generally* Michael S. Jacobs, *The New Sophistication in Antitrust*, 79 MINN. L. REV. 1, 8-12 (1994).

28. U.S. DEP'T. OF JUSTICE & FEDERAL TRADE COMM'N, MERGER GUIDELINES—1992, *reprinted in* 4 Trade Reg. Rep. (CCH) ¶ 13,104, at 20,569 (Apr. 17, 1997).

29. *Id.* § 1.51(c) (Certain mergers in "highly concentrated markets" presumptively create or enhance market power, unless the presumption is overcome "by a showing that the [other, structural market] factors make it unlikely" that those consequences will result.); *see also id.* § 1.51(b) (Certain mergers in "moderately concentrated markets post merger potentially raise significant competitive concerns depending on [other market] factors.").

30. See, for example, *United States v. Baker Hughes, Inc.*, 908 F.2d 981, 992 (D.C. Cir. 1990), where Judge Thomas declared as follows:

> Imposing a heavy burden of production on a [merger] defendant would be particularly anomalous where, as here, it is easy to establish a prima facie case. The government, after all, can carry its initial burden of production simply by presenting market concentration statistics. To allow the government virtually to rest its case at that point, leaving the defendant to prove the core of the dispute, would grossly inflate the role of statistics in actions brought under section 7.

31. *See, e.g.*, Jacobs, *supra* note 27, at 8 (arguing that "quixotic searches" for increasingly complex factors somehow thought germane to merger analysis are "inevitably inconclusive and all ultimately destructive of methodological simplicity and administrative efficiency").

32. For a discussion on the emerging use of innovation markets in merger analysis, see Jonathan B. Baker, *Fringe Firms and Incentives to Innovate*, 63 ANTITRUST L.J. 621 (1995); Richard J. Gilbert & Steven C. Sunshine, *Incorporating Dynamic Efficiency Concerns in Merger Analysis: The Use of Innovation Markets*, 63 ANTITRUST L.J. 569 (1995).

recent merger of two large pharmaceutical companies, requiring the post-merger firm to share certain research with its competitors.[33] During the same period, the FTC has developed what appears to be a new theory of market power[34]—unilateral competitive effects—to explain how certain mergers between sellers of differentiated products can create significant pockets of market power, even if they fall short of producing undue concentration in the relevant market as a whole.[35] While these new methodologies may be hard to comprehend, they do serve to make one thing quite clear: antitrust analysis is not getting any simpler.

Second, it is clear that these new approaches and technical twists in merger analysis are inspired by theory, not fact. They arise for the most part from a distinct economic philosophy, the post-Chicago School of antitrust economics, that has emerged over the past decade as an important counterweight to the hitherto unquestioned primacy of the Chicago School's theories. I have described and analyzed this development elsewhere[36] and need not rehearse it again here, save to say that theory, not fact, lies at the heart of both schools and that both admit, as indeed they must, that their theories are largely lacking in empirical support.[37] The new approaches to merger analysis, like their antecedents, presume that the business world operates as theory suggests it should.

B. Merger Analysis Applied to Hospital Mergers

If merger law in general is somewhat (or mainly) unsettled, hospital merger law finds itself in an even greater state of confusion. There are good reasons for this predicament. First, though mergers had been providing grist for antitrust adjudication prior to the enactment of the first federal statutes in 1890, it was not until 1975 that the Supreme Court definitively announced that the work of professionals—lawyers in that case—was "trade or commerce" within the meaning of the Sherman Act and therefore subject to antitrust scrutiny.[38] It took

---

33. This was the merger by which Ciba-Geigy and Sandoz became Novartis. *In re* Ciba-Geigy Ltd. et al., File No. 961-0055 (F.T.C. Dec. 5, 1996) (accepting a proposed consent decree for public comment to settle allegations that the proposed merger between Ciba-Geigy and Sandoz violated Section 7 of the Clayton Act in the gene therapy research, corn herbicide, and flea control product markets); *see also* Stephen Moore, *Drug Giant Novartis Voices Optimism, Despite Strains of its $27 Billion Merger*, WALL ST. J., June 18, 1997, at A17.

34. Though it might be simply a new method of defining the relevant product market.

35. *See generally* Jonathan B. Baker, *Mergers Among Sellers of Differentiated Products*, ANTITRUST, Spring 1997, at 23; Gregory J. Werden, *Simulating Unilateral Competitive Effects from Differentiated Products Mergers*, ANTITRUST, Spring 1997, at 27. This methodology is particularly sensitive to estimates of factors, such as elasticities of demand and supply, that are widely acknowledged to be extremely difficult, if not impossible, to estimate accurately.

36. *See* Jacobs, *supra* note 3 and accompanying text.

37. *Id.* at 251-54.

38. Goldfarb v. Virginia State Bar, 421 U.S. 773, 788 (1975).

more than a decade before a hospital merger case made its way into the federal reports.[39] By that time, however, although the core presumption of merger law—that anticompetitive consequences flow from undue concentration—had proven itself remarkably robust if not perfectly explanatory, it was nevertheless derived from the merger experience of "traditional" industry and commerce conducted by for-profit firms in markets that were largely if not entirely unregulated. Courts could plausibly regard these "traditional" markets as workably competitive; and thus could plausibly presume that big mergers would disturb a competitive status quo.

The hospital markets of the mid-1980s seemed very different from the "traditional" markets that had informed the development of merger law. They were not competitive in the usual sense of the word, but were heavily regulated in almost all aspects of their operation and were populated largely by not-for-profit firms—indeed, different kinds of not-for-profit firms—whose goals and objectives were arguably not the same as those of for-profit companies.[40] The financial subsidies from which hospitals benefitted generated substantial over-capacity in beds and equipment. When mergers began to occur in these markets, one might have argued that the presumptions animating traditional merger law did not hold. Because the hospital markets were distorted by virtue of having never operated competitively, the traditional presumption about the anti-competitive effects of increased concentration might not prove true. One might have also argued, equally reasonably, that the prevalence of not-for-profit firms in hospital markets altered the traditional presumption, because those firms might not be the unabashed profit-maximizers that populate the for-profit sector.

However, these arguments were not made when hospital merger law was a-borning. There were no solid empirical studies supporting these arguments; indeed, despite the judiciary's occasional plea for better evidence,[41] data-gathering efforts were slow to form. Instead, in the absence of hard evidence to the contrary, courts presumed for the most part that hospital markets were essentially similar to all others and thus applied to hospital mergers the same analytical approach used to judge mergers in other markets.[42] This early, reflexive approach to hospital mergers was not without its critics, in fact, powerful lobbying groups succeeded in modifying it on the federal level and, in many instances, eliminating it on the state level.[43] In the federal courts, however,

---

39. Hospital Corp. of Am. v. FTC, 807 F.2d 1381 (7th Cir. 1986).

40. *See, e.g.*, United States v. Mercy Health Servs., 902 F. Supp. 968, 973 (N.D. Iowa 1995) ("Traditionally, hospitals competed on the basis of amenities and perceptions of quality. Only in the last ten to fifteen years have hospitals begun to compete on the basis of price."), *vacated as moot*, 107 F.3d 632 (8th Cir. 1997).

41. *See, e.g.*, United States v. Rockford Mem'l Corp., 898 F.2d 1278 (7th Cir.), *cert. denied*, 498 U.S. 920 (1990).

42. *See, e.g.*, *id.*; FTC v. University Health, Inc., 938 F.2d 1206 (11th Cir. 1991).

43. U.S. Dep't of Justice & Federal Trade Comm'n, 1996 Statement of Enforcement Policy on Mergers Among Hospitals, *reprinted in* 4 Trade Reg. Rep. (CCH) ¶ 13,153 at 20,801 (Sept. 5, 1996) (outlining an "antitrust safety zone" in which a merger "between two general acute-care

with some notable exceptions, this approach has continued to hold sway though with subtle differences of opinion: for some, a hospital's not-for-profit status is irrelevant, for others it is dispositive;[44] also, for some, the relevant geographic markets are expanding, while for others they remain local.[45] In the grand scheme of things, though, these are small differences, because most courts continue to embrace the traditional notion—unsupported by empirical evidence—that hospital markets are no different fundamentally from other forms of business activity.

III. THE *BUTTERWORTH* CASE

A. The District Court Opinion

In *FTC v. Butterworth Health Corp.*,[46] the FTC sought to enjoin the merger of the two largest general acute care hospitals in Grand Rapids, Michigan, Butterworth Health Corporation and Blodgett Memorial Medical Center. Both were not-for-profit corporations, prosperous and well-managed.[47] Their merger was prompted by Blodgett's interest in constructing a $187 million replacement facility that would free it from the confines of an apparently undesirable location.[48] Shortly after Blodgett's relocation proposal failed to receive the approval of a county commission, Blodgett and Butterworth agreed to merge "to avoid substantial capital expenditures and achieve significant operating

---

hospitals where one of the hospitals (1) has an average of fewer than 100 licensed beds over the three most recent years, and (2) has an average daily inpatient census of fewer than 40 patients over the three most recent years" will not be challenged "about extraordinary circumstances").

44. *Compare Rockford Mem'l Corp.*, 898 F.2d at 1285 ("We are aware of no evidence—and the defendants present none, only argument—that nonprofit suppliers of goods or services are more likely to compete vigorously than other profit-making suppliers."), *and University Health, Inc.*, 938 F.2d at 1224 ("[T]he nonprofit status of the acquiring firm will not, by itself, help a defendant overcome the presumption of illegality that arises from the government's prima facie case."), *with* United States v. Carilion Health Sys., 707 F. Supp. 840, 849 (W.D. Va.) (concluding that the defendant's "nonprofit status weighs in favor of their mergers being reasonable"), *aff'd on other grounds*, 892 F.2d 1042 (4th Cir. 1989).

45. *Compare Mercy Health Servs.*, 902 F. Supp. at 978 ("It is not sufficient to take a snapshot of the current situation and define the relevant geographic market to be synonymous with the current service areas of the defendant hospitals."), *and* FTC v. Freeman Hosp., 911 F. Supp. 1213, 1221 (W.D. Mo.) (accepting the principle of geographic proximity, which identifies alternative sources consumers can seek out without further travel, to determine the relevant geographic market), *aff'd*, 69 F.3d 260 (8th Cir. 1995), *with Rockford Mem'l Corp.*, 898 F.2d at 1284-85 (determining that relevant geographic market was the hospitals' service area not the ten-county area proposed by the defendants).

46. 946 F. Supp. 1285 (W.D. Mich. 1996).

47. *Id.* at 1288.

48. *Id.* at 1288-89.

efficiencies."[49] The FTC challenged the merger, contending that it would unduly concentrate the relevant product and geographic markets.[50]

Agreeing with the FTC's proposed product and geographic market definitions, the district court identified two separate product markets, each with its own geographic market.[51] One product market consisted of general acute care inpatient hospital services—the garden variety product market in recent hospital merger cases;[52] the other market consisted of primary care inpatient hospital services, a more narrowly drawn market hitherto unrecognized in merger litigation.[53] For the first, broader product market, the court defined the relevant geographic market as Greater Kent County, Michigan, an area encompassing Grand Rapids and parts of seven adjoining counties within a thirty-mile radius.[54] Within that market, the court found that four Grand Rapids hospitals and five small rural hospitals competed for general acute care inpatients.[55] Regarding the smaller market for primary care services, the court identified the geographic market as the "immediate Grand Rapids area;" the only competitors in that market were the four Grand Rapids hospitals.[56]

Having identified the markets, the court then examined the extent to which the proposed merger would concentrate them. In that regard, the court found that the post-merger hospital corporation could account for as much as sixty-five percent of the general acute care inpatient market and seventy percent of the primary care inpatient market.[57] Additional evidence strongly suggested that other factors would exacerbate the concentrative effects of the merger and place the relevant markets under the control of a truly dominant firm: the two other Grand Rapids hospitals were found to offer a more limited range of services and to have a reputation for providing a lower quality of care;[58] entry barriers were significant;[59] and the merger partners had announced their intention to reduce certain managed care discounts post-merger and institute "standard managed care rates."[60] Persuaded by this evidence that "the merged entity would have substantial market power in two relevant markets," the court found that the FTC

---

49. *Id.* at 1289.

50. *Id.* at 1288-90.

51. *Id.* at 1290.

52. *See* FTC v. Freeman Hosp., 911 F. Supp. 1213, 1217 (W.D. Mo.), *aff'd* 69 F.3d 260 (8th Cir. 1995); United States v. Mercy Health Servs., 902 F. Supp. 968, 971 (N.D. Iowa 1995); FTC v. University Health, Inc., 938 F.2d 1206, 1210 (11th Cir. 1991).

53. These services include normal childbirth, gynecology, pediatrics, general medicine and surgery. *See Butterworth*, 946 F. Supp. at 1291.

54. *Id.* at 1290.

55. *Id.* at 1291.

56. *Id.* at 1293.

57. *Id.* at 1294.

58. *Id.* at 1298.

59. *Id.* at 1297-98.

60. *Id.* at 1299.

had established a prima facie case.[61]

Despite this evidence, however, the court refused to enjoin the merger.[62] Rather, it concluded on the basis of various pieces of more or less "empirical" evidence that defendants had rebutted the government's prima facie case by demonstrating that the admitted increase in market concentration would not harm consumers.[63] It used "facts," in other words, to overcome the presumption that dominant firms in highly concentrated markets will raise price to supra-competitive levels.[64]

The most important of these facts was the not-for-profit status of the merging hospitals. Despite the well-publicized unwillingness of most other courts to distinguish between for-profit and not-for-profit hospitals for purposes of merger analysis,[65] the *Butterworth* court found the two fundamentally different. Primarily on the basis of one recent study, which was fortified by some testimony and the judge's own "casual" empiricism, the court determined that not-for-profit hospitals in highly concentrated markets follow a significantly different pricing pattern from that of their for-profit counterparts. In particular, it found that while dominant for-profit hospitals set prices substantially above competitive levels, not-for-profit hospitals tend to *lower* prices when they attain market dominance.[66] This one study, the court implied, could overcome the force of merger law's basic presumption about the anticompetitive impact of high concentration.[67]

There were other pieces of "data" that also proved persuasive to the court. The court observed, for example, that the boards of the merging hospitals consisted of local business leaders who, it said, "have a direct stake in maintaining high quality, low cost hospital services."[68] It considered the testimony of the board chairmen, each of whom testified "convincingly" that the merger was motivated "by a common desire to lower health care costs and

---

61. *Id.* at 1302.

62. *Id.* at 1303.

63. *Id.* at 1302.

64. *Id.*

65. *See* FTC v. University Health, Inc., 938 F.2d 1206, 1225 (11th Cir. 1991); United States v. Rockford Mem'l Corp., 898 F.2d 1278, 1286 (7th Cir.), *cert. denied*, 498 U.S. 920 (1990); United States v. Mercy Health Servs., 902 F. Supp. 968, 989 (N.D. Iowa 1995); *see also In re* Hospital Corp. of Am., 106 F.T.C. 361, 502 (1985), *aff'd*, 807 F.2d 1381, 1393 (7th Cir. 1988). Some courts, however, have suggested that mergers between not-for-profit hospitals merit distinctive and more lenient treatment. *See* FTC v. Freeman Hosp., 911 F. Supp. 1213, 1226-28 (W.D. Mo.), *aff'd*, 69 F.3d 260 (8th Cir. 1995); United States v. Carilion Health Sys., 707 F. Supp. 840, 847-49 (W.D. Va.), *aff'd on other grounds*, 892 F.2d 1042 (4th Cir. 1989).

66. *See Butterworth*, 946 F. Supp. at 1296-97. The court based its conclusion on empirical studies by economists showing that high market concentration with dominant not-for-profit hospitals does not correlate positively with higher prices and may result in lower prices. *See* William J. Lynk, *Nonprofit Hospital Mergers and the Exercise of Market Power*, 38 J.L. & ECON. 437, 459 (1995).

67. *Butterworth*, 946 F. Supp. at 1297.

68. *Id.* at 1296.

improve the quality of care."[69] It toured the hospitals and came away "confirmed" in its belief that the boards had been "responsible stewards of the resources available to them."[70] The court was impressed by the defendants' offer to bind themselves contractually—through a "Community Commitment"—not to raise prices "or otherwise injure the community."[71] And it noted that within the relevant markets important buyers of health care—those presumably at risk of any monopoly pricing that the merger might permit—supported the merger.[72] These and other reasons[73] convinced the district court that though it might well produce undue concentration in the relevant markets, the proposed merger would probably not generate anticompetitive effects. In a per curiam opinion, the Court of Appeals for the Sixth Circuit affirmed, concluding that the district court had not abused its discretion.[74]

B. Empiricism Examined: the Lynk Study, the Role of the Non-Profit Board, and Community Commitment

The most "empirical" portion of the proof in *Butterworth* was a study published in 1995 by William Lynk in the well-respected *Journal of Law and Economics*.[75] Despite the FTC's objection to the study's methodology, the court relied heavily on its conclusions.[76] The study examined the post-merger pricing behavior of California hospitals that became dominant in their markets through merger. Looking at these hospitals by category,[77] the court found that compared to for-profit hospitals, certain private not-for-profit hospitals in the study sample "have a significantly lower association between higher market shares and higher prices, and on balance increased nonprofit market share is associated with lower,

---

69. *Id.* at 1297.

70. *Id.* at 1299.

71. *Id.* at 1298. The boards promised to (1) freeze list prices or charges, (2) freeze prices to managed care plans at pre-merger levels, (3) limit profit margins, (4) provide for the underserved and medically needy, and (5) establish a governing board for the merged entity that included community representation reflective of the diversity of West Michigan. *Id.* at 1304-06.

72. *Id.* at 1299.

73. The court also found that the merger "would result in significant efficiencies, in the form of capital expenditure avoidance and operating efficiencies, totaling in excess of $100 million." *Id.* at 1301. It made this finding after comparing a comprehensive study prepared by defendants with the FTC's expert's critical analysis, and after making its own tour of the defendants' facilities and concluding that in the absence of the merger, Blodgett would certainly proceed with its relocation plans, and Butterworth would respond by renovating and upgrading its own facilities, setting off a "medical arms race" that would harm consumers. *Id.*

74. FTC v. Butterworth Health Corp., 121 F.3d 708 (6th Cir. 1997).

75. Lynk, *supra* note 66, at 442.

76. *Butterworth*, 946 F. Supp. at 1296.

77. The categories were for-profit hospitals, private not-for-profit hospitals, and government-run hospitals. Lynk, *supra* note 66, at 442.

not higher, prices."[78]

The court also considered a second study, described as a "replication" of the first, that was prepared by Lynk expressly for the underlying litigation and focused strictly on the pricing behavior of Michigan hospitals. This study concluded that in Michigan, as in California, mergers between not-for-profit hospitals that resulted in highly concentrated markets were associated with lower post-merger prices.[79] After evaluating this data, the FTC's expert economist "obtained similar results concerning the relationship between market concentration and pricing levels,"[80] and though the two experts disputed the likely cause of the findings,[81] they agreed that "high market concentration among nonprofit hospitals does not correlate positively with higher prices."[82]

The court placed great weight on these studies. By showing that high market concentration does not lead not-for-profits to raise prices, these "unexpected empirical findings" cast doubt, in the court's view, on the traditional presumption of merger law that a significant increase in market concentration will lead to higher prices.[83] This doubt about the traditional presumption, coupled with evidence of anticipated cost savings from the merger and the defendants' promise to freeze prices for three years following the merger, led the court to presume in turn that these particular defendants would not raise their prices post-merger.[84] The court thus allocated to the government the burden of overcoming that presumption through evidence of post-merger anticompetitive effects.[85] Such evidence is very difficult to muster, since it is necessarily speculative, and the government could not do it in this case. It was reduced to questioning the precise

---

78. *Butterworth*, 946 F. Supp. at 1295 (quoting Lynk, *supra* note 66, at 459). This result was reported for what the author labeled "consumer co-op" hospitals, those where the hospital's consumers control its general policies—through representation on the board and "perhaps through other means." Lynk, *supra* note 66, at 441. The study found that other not-for-profit hospitals—the "government (usually county) hospital"—perhaps because their income could be made available to fund other undertakings of the owner—priced in a manner roughly comparable to for-profit hospitals. *Id.* at 459. Lynk postulated that this phenomenon may be explained by the potential of mergers to create "economic efficiencies" and by the possibility that "a nonprofit hospital organization whose only function is the provision of hospital services to a well-defined population, and whose governing board effectively represents the same population, looks—and probably acts—more like a consumer cooperative than a creator of monopoly resource misallocation." *Id.* at 458.

79. *Butterworth*, 946 F. Supp. at 1295.

80. *Id.*

81. *Id.* Lynk argued that the lower prices resulted from merger-specific efficiencies, "through the consolidation of clinical services and other means." *Id.* (quoting Lynk, *supra* note 66, at 458). Leffler, the government's expert, contended it might have to do less with efficiencies and more with the "ruralness"—and thus lower costs—of concentrated hospital markets. *Id.*

82. *Id.*

83. *Id.*

84. *Id.* at 1297-98.

85. *Id.* at 1298.

amount of the anticipated cost-savings and the strength of the hospitals' promise to freeze prices; but since the government's vision of the future was no more plausible than the defendants,' the merger was allowed to proceed.[86]

In an important sense then, the empirical studies proved outcome-determinative. By persuading the court to drop the presumption that higher prices follow high concentration, they effectively caused the burden of proof to shift from the defendants to the government, and caused the nature of the proof about anti-competitive effects to change from theoretical to "factual." Since this burden was—not surprisingly—impossible to carry, the government lost.

In *Butterworth*, the government challenged the methodology used in both Lynk studies.[87] As others in this symposium have persuasively argued, empirical studies can often go wrong or confuse more than they illuminate, for a wide variety of reasons.[88] In my view, however, the *Butterworth* court was right to employ and rely upon the Lynk studies. In the first place, although antitrust in general—merger law no less—depends upon presumptions for ease of administration, predictability and the like, courts have repeatedly emphasized that these presumptions serve useful ends only when they reflect facts. The Supreme Court issues periodic reminders about the limits of legal presumptions in antitrust law generally,[89] and at least one Court of Appeals has rued the absence of empirical data regarding the workings of the hospital markets and issued a call for more research into this area.[90]

The Lynk studies purport to show important facts that bear on the competitive nature of not-for-profit hospitals. This is important information about market structure and deserves consideration; indeed, given our relatively undeveloped understanding about the nature of competition between not-for-profit hospitals, it would be foolish to disregard these studies. If legal presumptions are meant to be rooted in fact, the existence of new, salient facts has a direct bearing on the continuing validity of old presumptions. As new facts help to displace old presumptions, the law may seem rudderless for a while, too particularistic, lacking in guidance or intelligibility. But this is a reasonable price to pay for new, sensible presumptions. As more new facts emerge, new presumptions will form around them; presumptions in general are useful, but no particular presumption has a right to eternal life.

This is not to say that every empirical study is a good one or that all are

---

86. *Id.* at 1302-03.

87. *Id.* at 1296.

88. *See, e.g.*, William M. Sage, *Judicial Opinions Involving Health Insurance Coverage: Trompe L'Oeil or Window on the World?*, 31 IND. L. REV. 49, 61-68 (1998).

89. *See, e.g.*, Eastman Kodak Co. v. Image Technical Servs., Inc., 504 U.S. 451, 466 (1992) ("Legal presumptions that rest on formalistic distinctions rather than actual market realities are generally disfavored in antitrust law. . . . In determining the existence of market power . . . this Court has examined closely the economic reality of the market at issue.").

90. *See* United States v. Rockford Mem'l Corp., 898 F.2d 1278, 1286 (7th Cir.), *cert. denied*, 498 U.S. 920 (1990) ("It is regrettable that antitrust cases are decided on the basis of theoretical guesses as to what particular market-structure characteristics portend for competition.").

equally deserving of judicial consideration. Empirical studies can be flawed, inconclusive, untimely, and otherwise problematic. But they can also prove very valuable. As a class they are arguably relevant to the work of antitrust courts, each subject to attack by opposing counsel and experts on traditional grounds of evidentiary sufficiency. Indeed, perhaps in the fullness of time and with the assistance of the academic community, courts can develop comprehensible rules of easy application for assessing the validity of empirical methodologies. But until then, cross-examination and the opposing testimony of expert witnesses can help expose flawed studies. Lynk's work survived these challenges. The *Butterworth* court's use of that work advances the cause of antitrust law by moving it in the direction of greater accuracy.

This brings us to the next piece of collective data. This includes the *Butterworth* court's observation that the boards of the merging hospitals consisted of local business leaders who, it said, "have a direct stake in maintaining high quality, low cost hospital services."[91] It considered the testimony of the board chairmen, each of whom testified "convincingly" that the merger was motivated "by a common desire to lower health care costs and improve the quality of care."[92] It was impressed by defendants' offer to bind themselves through a "Community Commitment" not to raise prices "or otherwise injure the community."[93] It noted that important buyers of health care—those who would presumably be at risk for any monopoly pricing that the merger might permit—supported the merger.[94] The court rejected the government's claim that the managed care discounting foreclosed by the merger would harm consumers generally, on the basis of evidence—not clearly specified—that consumers in managed care were outnumbered by those covered by more traditional forms of health insurance.[95]

There are many ways to evaluate the court's use of this data. The court could be regarded as having made a factual finding that dominant not-for-profit hospitals whose board members are drawn from the community are structurally disinclined to raise prices to the people that they serve. The same William Lynk who authored the pricing studies discussed above has hypothesized—in non-empirical fashion—that the "consumer cooperative" not-for-profit hospital lacks an economic incentive to raise prices to its members.[96] But the court did not express such a finding, and, although theory suggests that corporate structure

---

91. *Butterworth*, 946 F. Supp. at 1296.

92. *Id.* at 1297.

93. *Id.* at 1298. The boards promised to (1) freeze list prices or charges, (2) freeze prices to managed care plans at pre-merger levels, (3) limit profit margins, (4) provide for the underserved and medically needy and (5) establish a governing board for the merged entity with community representation that reflected the diversity of West Michigan. *Id.* at 1298, 1304-06.

94. *Id.* at 1299.

95. *Id.* at 1298-99.

96. *See* William J. Lynk, *Property Rights and the Presumptions of Merger Analysis*, ANTITRUST BULL. 363, 377 (1994); *see also* Henry B. Hansmann, *The Role of Nonprofit Enterprise*, 89 YALE L.J. 835, 889 (1980).

might discourage certain kinds of not-for-profits from using market power to harm consumers, the court lacked the empirical warrant necessary for such a broad conclusion.

On the other hand, the *Butterworth* court could be seen as having found that these defendants—because of their governing structure and their Community Commitment to hold the line on price—would not raise price post-merger. I believe that the court did make such a finding, and I view it as problematic. The empirical question here is whether the Community Commitment (1) buttresses the Lynk study on post-merger pricing and (2) is judicially administrable in the event of breach. Regarding the first point, the Commitment seems to add little of substance to the broader empirical question: it is not a promise to lower prices—despite the acknowledged efficiencies that the merger will yield—but rather a promise not to raise them. Regarding the second point, by bringing the promise into its merger analysis, the court seems to be inviting future litigation of a regulatory nature about whether the promise has been breached by an unwarranted price increase. This kind of oversight might be workable if undertaken by the State Attorney General as part of a compromise to ensure the preservation of social and community service, but the federal courts might be hard pressed to undertake this task.

Finally, the court's attention to buyer attitudes towards the prospect of a post-merger price increase could be viewed as a useful attempt to employ a workable rule of thumb to the complex factual questions in merger analysis. In *Butterworth*, large local buyers of hospital services supported the merger.[97] Despite their exposure to post-merger monopoly overcharges (or lowering of quality, or service cutbacks), those who would be buying from the merged hospital not only failed to oppose the merger but actually supported it. This is important data. Rational buyers, fearing future harm in the form of higher prices, would strenuously argue against a merger creating a profit-maximizing monopolist. Since we have no reason to doubt the rationality of large buyers in Grand Rapids, their support of the merger must mean that they think that they will be better off after the merger. How can this be? In all likelihood, they share the hospitals' view regarding the efficiencies flowing from the merger and trust the hospitals to pass on efficiency-related cost-savings in the form of lower prices.

Buyer reaction thus provides a good test—a useful rule of thumb—of the merger's likely effects. Buyers are the victims of anticompetitive mergers, forced to pay supra-competitive prices. Merger law exists, largely if not exclusively, to protect them. If they disdain "protection" and welcome a merger, antitrust has no reason to override their collective will, provided that the buyers in question are (1) rational, which we must assume in the absence of good evidence to the contrary, and (2) sufficiently representative of all buyers in the market, or (3) sufficiently numerous so that we can infer representativeness. In *Butterworth*, the court simply recited that the buyers were "important," without indicating precisely what it meant by the term. Presumably the buyers were large

---

97. *Butterworth*, 946 F. Supp. at 1299.

and representative—the government did not argue otherwise—but it would have been preferable for analytical purposes to know somewhat more about them. Nevertheless, the court made good use, in my opinion, of a valuable empirical test inexpensive to develop and easy to apply: for the best assessment of a merger's likely effects, ask a fair sampling of buyers. Antitrust needs more tests like this.

CONCLUSION

It is difficult to determine when the state of empirical research has advanced sufficiently to serve as a basis for changing the law. No one would claim, I should imagine, that courts should be oblivious to new scientific findings[98] or that, once adopted, legal presumptions whose factual predicates no longer hold true should continue in force. Nor, however, would anyone advocate that courts adopt every new study offered for their consideration, changing the law to suit the most recent set of findings. They must strike a balance respecting both the teachings of science and the need for stable doctrine, a balance arguably more necessary in antitrust—given the size of the stakes and the need for predictability—than in most areas of law.

The *Butterworth* opinion has its problems, some of them substantial.[99] But it makes a good attempt at striking the necessary balance. It will no doubt receive its share of criticism, but the critics will be talking details: were Lynk's studies "good enough;" should there be rules for the admissibility of such studies; were the buyers whose reactions were gauged adequately representative of buyers as a whole? These are good questions, but they hardly detract from the legitimacy of reconsidering old presumptions in the light of new evidence and attempting to develop workable rules of antitrust analysis.

---

98. *See* Daubert v. Merrill Dow Pharms., 509 U.S. 579 (1993) (providing a more formal means of incorporating even marginal scientific evidence into the trial process); *see also* Petruzzi's IGA Supermkts., Inc. v. Darling-Delaware Co., 998 F.2d 1224, 1241 (3d Cir. 1993) (antitrust case raising *Daubert* issues with regard to expert economic testimony).

99. One of the biggest problems is that by relying on the current preferences of large buyers, most of whom do not purchase health insurance from managed care organizations, the court effectively freezes current preferences in place, handicapping the future growth of managed care in Grand Rapids by eliminating the possibility that managed care organizations could obtain low hospital prices by pitting one competing hospital against the other. Perhaps, however, buyers foresee this possibility but do not welcome it because they perceive that heightened price competition between hospitals does not necessarily benefit the community at large.

Indiana Law Review

Volume 31 **1998** **Number 1**

NOTES

THE STATE OF COPYRIGHT PROTECTION FOR ELECTRONIC DATABASES BEYOND *PROCD V. ZEIDENBERG*: ARE SHRINKWRAP LICENSES A VIABLE ALTERNATIVE FOR DATABASE PROTECTION?

JENNETT M. HILL*

INTRODUCTION

J-CAP Directories, Inc., specializing in computer disc technology, invests $10 million researching and developing a comprehensive telephone listings database. The marketing department finds great demand for such a listing in the form of an electronic compilation. So, researchers compile 95 million listings and incorporate them into an electronic database. The listings include residential and commercial names, addresses and telephone numbers. The business listings contain pertinent industry data to enable marketing companies to target their customers. Software and database technicians develop a state-of-the-art database and access software while data entry personnel begin loading the 95 million records into the database. This database application[1] enables the computer user to search using any number of fields: name, town, state, zip-code, industry code and telephone area code. The company stores the database application on CD-ROM[2] discs and packages the product, complete with a "how-to" manual, in an appealing box encased in cellophane. The box displays a warning that the purchase of this product is subject to a restrictive licensing agreement. In fact, the creative software engineers included a licensing agreement screen that appears each time the software is executed, forcing the user to respond by pressing "Enter."

The CD-ROM product is a tremendous success, until an entrepreneurial computer programmer purchases a copy of the database application. The entrepreneur downloads the CD-ROM data onto his computer and uploads J-

---

* J.D. Candidate, 1998, Indiana University School of Law—Indianapolis; B.A., 1985, DePauw University, Greencastle, Indiana. The author would like to thank Professor Kenneth D. Crews for his guidance in the initial draft of this Note. The author would also like to give special thanks to her husband, Alan P. Hill, for his love, encouragement and support that made this Note—and law school—possible.

1. A database application, for purposes of this Note, includes the database (compilation of data) and the software used to access the data.

2. CD-ROM: Compact disk-read only memory.

CAP's sophisticated database onto an Internet host computer. The computer wizard develops his own software to access the downloaded database across the Internet. He then sells access to the Internet database for a fee. The entrepreneur accomplishes this in less than three months, investing only $50,000.

J-CAP discovers its million dollar investment on the Internet and files suit alleging copyright infringement and a breach of the license agreement. To the company's surprise, the court finds no database protection under copyright law and holds the license agreement unenforceable as a matter of contract law! In essence, J-CAP Directories spent millions on a product it could not protect from subsequent developers. This hypothetical demonstrates the commercial inequities of database protection under copyright law. Within the last five years the courts have eroded and then propped up electronic database protection under both copyright and contract law. This Note analyzes the recent developments of copyright protection as it pertains to electronic databases.

Lacking sufficient copyright protection, many companies and state governments have turned to "shrinkwrap licenses,"[3] as an alternative to database protection. However, this raises a fundamental question: should other approaches to electronic database protection prevail in light of weak and unclear copyright protection given the specific preemption clause in § 301 of the Copyright Act?[4] This Note examines these approaches and submits that the United States legislature should evaluate the enactment of a *sui generis* right for database protection to protect valuable economic and global resources.

Part I of this Note addresses the dilemma of electronic database protection through an analysis of two recent cases: *Feist Publications, Inc. v. Rural Telephone Service Co.*[5] and *ProCD, Inc. v. Zeidenberg.*[6] These cases highlight database protection issues under current copyright law. Part II of this Note discusses the evolution of the Copyright Act and its impact on computerized materials. This section discusses the preemption clause of the Act, and the role it plays in determining the validity of shrinkwrap licenses. This section further examines Congress' intent in enacting its 1976 revision of the Act and whether their policy and goals continue to guide the judiciary. Part III summarizes the state of database protection in view of the preceding cases and copyright law. Due to uncertain copyright protection for electronic databases, companies developed other methods of protection. Part IV evaluates these developments, specifically, shrink-wrap licenses as successfully secured in *ProCD*;[7] the recent draft of the Uniform Commercial Code, which includes a new licensing section; and the international community's database protection proposals issued to the

---

3. The term "shrinkwrap licenses" evolved from using a cellophane or plastic wrapping to package computer applications. Companies often enclosed a license agreement and referred to a brief provision of the agreement which presumably became effective once the purchaser opened the cellophane (or shrinkwrap) or executed the program. *See infra* note 154 and accompanying text.

4. 17 U.S.C. § 301 (1994).

5. 499 U.S. 340 (1991).

6. 86 F.3d 1447 (7th Cir. 1996).

7. *Id.*

World Intellectual Property Organization. This section also discusses the United States' response to international pressures regarding database protection. Finally, Part V analyzes methods for future database protection.

I. CASE LAW IMPACT ON DATABASE PROTECTION

Case law developments have weakened database protection under copyright law. In 1991, the Supreme Court eviscerated a popular theory upon which many courts and database developers relied—the "sweat of the brow" theory.[8] With reduced copyright protection for databases, the technology industry turned to contract law by adopting the "shrinkwrap license" as their preferred method of protection. Some courts have declared shrinkwrap licenses unenforceable as a matter of contract law[9] and federal preemption laws.[10] In 1996, however, the Seventh Circuit held that a shrinkwrap license, which limited the use of a database application and secured exclusive rights in the database, was enforceable under contract law and was not preempted by the Copyright Act.[11]

A. Feist*: The Key to Protection is Originality*

In *Feist*, the Supreme Court held that "white pages" in a phone book were not subject to copyright protection.[12] Rural Telephone, a public utility company providing telephone service to communities in northwest Kansas, alleged that Feist Publications infringed on its copyright by publishing Rural's listings in Feist's area-wide directory.[13] To compile white pages listings, Feist solicited area telephone publishing companies requesting licensing permission and offering to pay for the right to publish the listings in its directory.[14] All companies agreed to grant a license except Rural.[15] To avoid a large gap in its listings, Feist published 1,309 of Rural's listings and added the individuals' street addresses to its directory.[16] The Supreme Court granted *certiorari* to determine "whether the copyright in Rural's directory protects the names, towns, and telephone numbers copied by Feist." [17]

The Court focused on the requirement of "originality" to secure copyright protection and conceded that the threshold requirement for establishing

---

8. The "sweat of the brow," or "industrious collection", theory protected a database when a significant amount of work was invested in compiling facts. *See Feist*, 499 U.S. at 352.

9. *See, e.g.*, Step-Saver Data Sys., Inc. v. Wyse Tech., Inc., 939 F.2d 91 (3d Cir. 1991); Arizona Retail Sys., Inc. v. Software Link, Inc., 831 F. Supp. 759 (D. Ariz. 1993).

10. *See, e.g.*, Vault Corp. v. Quaid Software Ltd., 847 F.2d 255 (5th Cir. 1988).

11. *ProCD*, 86 F.3d at 1455.

12. *Feist*, 499 U.S. at 362-63.

13. *Id.* at 344.

14. *Id.* at 343.

15. *Id.*

16. *Id.* at 344. Rural's directory contained approximately 7, 700 total listings. *Id.* at 342.

17. *Id.* at 344.

originality was low.[18] Although individual facts are not copyrightable, factual "compilations"[19] pass the requisite originality test because the author must select and arrange the facts.[20] The Court found an explicit originality requirement in the Copyright Act[21] as well as an implicit requirement of originality in the Copyright Clause of the Constitution.[22]

In defining originality, *Feist* denounced the "sweat of the brow" or "industrial collection" theories relied upon by lower courts to protect compilations. The court found that the purpose of copyright was to motivate authors to create works and not reward them based solely on industrious efforts.[23] "[T]he 1976 revisions to the Copyright Act leave no doubt that originality, not 'sweat of the brow,' is the touchstone of copyright protection in directories and other fact-based works."[24]

The Court in *Feist* upheld a compiler's "selection and arrangement" as a method of demonstrating originality in a compilation.[25] Thus, the selection and

---

18. "The *sine qua non* of copyright is originality." *Id.* at 345. "To be sure, the requisite level of creativity is extremely low; even a slight amount will suffice." *Id.*

19. "A 'compilation' is a work formed by the collection and assembling of preexisting materials or of data that are selected, coordinated, or arranged in such a way that the resulting work as a whole constitutes an original work of authorship." 17 U.S.C. § 101 (1994).

20. *Feist*, 499 U.S. at 348. "[C]hoices as to selection and arrangement, so long as they are made independently by the compiler and entail a minimal degree of creativity, are sufficiently original that Congress may protect such compilations through the copyright laws." *Id.* (citing 1 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT §§ 2.11[D], 3.03 (1990)).

21. *Id.* at 355 (citations omitted). The Court found the originality requirement in the phrase "original works of authorship" in § 102(a) of the Copyright Act. *Id.* "The two fundamental criteria of copyright protection [are] originality and fixation in tangible form . . . . The phrase 'original works of authorship,' which is purposely left undefined, is intended to incorporate without change the standard of originality established by the courts under the present [1909] copyright statute." *Id.* (quoting H.R. REP. No. 94-1476, at 51 (1976); S. REP. NO. 94-473, at 50 (1975), *reprinted in* 1976, U.S.C.C.A.N. 5659, 5664).

22. *Id.* at 351. The Court explained that "originality is a constitutionally mandated prerequisite for copyright protection." *Id.* The Copyright Clause authorizes Congress to "secur[e] for limited Times to Authors . . . the exclusive Right to their respective Writings." U.S. CONST. art. I., § 8, cl. 8. [hereinafter Copyright Clause]. Interestingly, this is the first time since 1879 that the Court has held a work unworthy of copyright protection on Constitutional grounds. Ralph Oman, *Reflections on the Changing Shape of Database Protection*, 40 FED. B. NEWS & J. 232, 237 n.18 (1993). In the 1800's, the Supreme Court decided two landmark cases that addressed the constitutional scope of works: *Burrow-Giles Lithographic Co. v. Sarony*, 111 U.S. 53 (1884); and *The Trade-Mark Cases*, 100 U.S. 82 (1879).

23. *Feist*, 499 U.S. at 349. "The primary objective of copyright is not to reward the labor of authors, but '[t]o promote the Progress of Science and useful Arts.'" *Id.* (quoting U.S. CONST. art. I., § 8, cl. 8).

24. *Id.* at 359-60.

25. *Id.* at 350. "[O]nly the compiler's selection and arrangement may be protected; the raw facts may be copied at will." *Id.*

arrangement of the factual compilation underlie an author's claim to originality. In *Feist*, however, the selection and arrangement "lack[ed] the modicum of creativity necessary to transform mere selection into copyrightable expression. Rural expended sufficient effort to make the white pages directory *useful*, but insufficient creativity to make it *original*."[26] Although the Court reminded us that the originality requirement is rooted in the Constitution, it did not ignore the statutory requirements of originality either:

> The [1976 Copyright Act] revisions explain with painstaking clarity that copyright requires originality, § 102(a); that facts are never original, § 102(b); that the copyright in a compilation does not extend to the facts it contains, § 103(b); and that a compilation is copyrightable only to the extent that it features an original selection, coordination, or arrangement, § 101.[27]

The Court was unwilling to stretch the concept of authorship to protect an author's labor.[28] Instead, only "those products that evince intellectual authorship" are protected under copyright law.[29] Consequently, *Feist* has diminished copyright protection for databases. Although the selection and arrangement theory remains, it merely prevents copying the format of the data, but does not protect the tedious, expensive and comprehensive collection of the data itself.[30]

The selection and arrangement theory has been applied in subsequent cases.[31] For example, the Second Circuit found that a publisher's valuation for used vehicles (the "Red Book") was protected as a compilation under *Feist*.[32]

---

26. *Id.* at 362-63 (emphasis added).

27. *Id.* at 360.

28. Oman, *supra* note 22, at 234.

29. *Id.*

30. *See generally* John F. Hayden, Recent Development, *Copyright Protection of Computer Databases After Feist*, 5 HARV. J.L. & TECH. 215, 236 (1991) (selection and arrangement "fails to protect the database's main contribution"—the collection of the data and database development); Jessica Litman, *After Feist*, 17 U. DAYTON L. REV. 607, 609 (1992) (author concedes that "computer databases are surely sufficiently original to merit copyright protection").

31. *See e.g.*, CCC Info. Serv., Inc. v. Maclean Hunter Market Reports, Inc., 44 F.3d 61 (2d Cir. 1994), *cert. denied*, 116 S. Ct. 72 (1995); BellSouth Adver. & Publ'g Corp. v. Donnelley Info. Publ'g, Inc., 999 F.2d 1436 (11th Cir. 1993), *cert. denied*, 510 U.S. 1101 (1994) (a telephone publisher's competitor was permitted to copy elements of the publisher's compilation (name, address, etc.) without infringement because the competitor did not copy the selection, coordination or arrangement of the data); Victor Lalli Enter. v. Big Red Apples, Inc., 936 F.2d 671 (2d Cir. 1991) (finding that a compilation of horse racing statistics formatted in a static fashion lacked sufficient selection and arrangement to substantiate an original work); Montgomery County Ass'n of Realtors, Inc. v. Realty Photo Master Corp., 878 F. Supp. 804 (D. Md. 1995), *aff'd*, 91 F.3d 132 (4th Cir. 1996).

32. *Maclean*, 44 F.3d at 67. "Compilations that devise new and useful selections and arrangements of information unquestionably contribute to public knowledge by providing cheaper,

"The fact that the arrangement of data responds *logically* to the needs of the market for which the compilation was prepared does not negate originality."[33] In 1995, a Maryland court found an electronic database of real estate listings accompanied by photographs met the selection and arrangement originality requirement and granted the creator copyright protection.[34] Nonetheless, many companies are unsure of the stability of copyright protection for databases, and have therefore resorted to using contracts to protect their works.

B. ProCD*: A Viable approach to Database Protection?*

ProCD invested ten million dollars to compile a comprehensive directory comprising 3,000 telephone directories in one computer database.[35] ProCD packaged this database application and sold the product as "SelectPhone" on CD-ROM.[36] On the outside of each box ProCD alerted purchasers that a license was enclosed in the database application.[37] This license, which limited use of the database application to non-commercial purposes, was not only printed in the manual, but also appeared on the computer screen each time the enduser executed the software.[38] Computer science student Matthew Zeidenberg purchased a consumer version of the database application and developed his own software to access the database.[39] He then placed the database on an Internet server and charged his clients to access the database through his company, Silken Mountain Web Services, Inc.[40] ProCD asserted that Zeidenberg went beyond the scope of the shrinkwrap license when he placed the database on the server and provided access across the Internet.[41]

Contrary to prior circuit court decisions,[42] the court in *ProCD* held that a

---

easier, and better organized access to information." *Id.* at 66.

33. *Id.* at 67 (emphasis added).

34. *Montgomery*, 878 F. Supp. at 810. The district court held that the computer database's arrangement was sufficient to sustain the originality requirement under *Feist*. The court concluded that the database developer added "marketing puffery" for each home which could not be categorized as factual, further supporting an original "presentation" and "arrangement." *Id.*

35. ProCD, Inc. v. Zeidenberg, 86 F.3d 1447, 1449 (7th Cir. 1996).

36. *Id.*

37. *Id.* at 1450.

38. *Id.* "Enduser" is used to refer to the person using the electronic database for searches. For example, a student accessing Westlaw's database to perform legal searches is an "enduser" for purposes of this Note.

39. *Id.*

40. *Id.*

41. *Id.*

42. Step-Saver Data Sys., Inc. v. Wyse Tech., Inc., 939 F.2d 91 (3d Cir. 1991) ("limited use license agreement" printed on package containing computer software did not become part of the parties' agreement and was therefore unenforceable); Vault Corp. v. Quaid Software Ltd., 847 F.2d 255 (5th Cir. 1988) (provisions in plaintiff's license agreement, which prohibited decompilation or disassembly of its program, was unenforceable); Arizona Retail Sys. Inc. v. Software Link, Inc.,

shrinkwrap license, imposing restrictions on an enduser, was enforceable.[43] In reaching its decision, the court relied upon contract law finding that "[a] vendor, as master of the offer, may invite acceptance by conduct, and may propose limitations on the kind of conduct that constitutes acceptance."[44] Further, "the U.C.C. consistently permits the parties to structure their relations so that the buyer has a chance to make a final decision after a detailed review."[45]

As to preemption, the court held that copyright law did not preempt private contracts as "courts usually read preemption clauses to leave private contracts unaffected. . . . [j]ust as [the copyright preemption clause] does not itself interfere with private transactions in intellectual property, so it does not prevent states from respecting those transactions."[46] Moreover, "whether a particular license is generous or restrictive, a simple two-party contract is not 'equivalent to any of the exclusive rights within the general scope of copyright' and therefore may be enforced."[47] Breaking precedent, the Seventh Circuit created a circuit split by enforcing shrinkwrap licenses in spite of copyright preemption.[48] In short, the Seventh Circuit endorsed the use of shrinkwrap licenses as a valid method for database protection.

II. HISTORICAL DEVELOPMENTS IN COMPUTER COPYRIGHT PROTECTION

A. Copyright Protection for Compilations

The Constitution grants Congress the power "[t]o promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors, the exclusive Right to their respective Writings and Discoveries."[49] The framers

---

831 F. Supp. 759 (D. Ariz. 1993) (buyer was aware of the terms of the license before purchasing the software because the buyer requested an evaluation diskette detailing similar terms). *See infra* note 125 and accompanying text.

43. *ProCD*, 86 F.3d at 1449.

44. *Id.* at 1452. The court, relying on several sections of the U.C.C. as adopted by Wisconsin, stated:

> A contract for sale of goods may be made in any manner sufficient to show agreement, including conduct by both parties which recognizes the existence of such a contract. . . . A buyer accepts goods under § 2-606(1)(b) when, after an opportunity to inspect, he fails to make an effective rejection under § 2-602(1). ProCD extended an opportunity to reject if a buyer should find the license terms unsatisfactory; Zeidenberg inspected the package, tried out the software, learned of the license, and did not reject the goods.

Id. at 1452-53 (citations omitted).

45. *Id.* at 1453.

46. *Id.* at 1454-55. The copyright preemption clause (§ 301(a)), in pertinent part, provides that "all legal or equitable rights that are equivalent to any of the exclusive rights within the general scope of copyright. . . are governed exclusively by this title." 17 U.S.C. § 301(a) (1994).

47. *ProCD*, 86 F.3d at 1455 (citing 17 U.S.C. § 301(a) (1994)).

48. *See infra* note 125 and accompanying text.

49. U.S. CONST. art. I, § 8, cl. 8.

intended copyright law to induce creative works which benefit the public.[50] However, protecting an author's creative rights may conflict with the public's right to access the work.[51] To balance these competing interests, Congress enacted the 1790 copyright statute.[52]

This first statute provided only limited protection [53] but, by 1909, copyright protection also included compilations of facts.[54] Today, the 1976 Copyright Act expressly covers nine categories of works including compilations.[55] "A 'compilation' results from a process of selecting, bringing together, organizing, and arranging previously existing material of all kinds, regardless of whether the individual items in the material have been or ever could have been subject of copyright."[56] The subject matter of a compilation, however, "extends only to the material contributed by the author of such work, as distinguished from the preexisting material employed in the work, and does not imply any exclusive right in the preexisting material."[57]

In response to emerging technologies, Congress formed a national commission to provide recommendations for copyright reform.[58] Based upon recommendations from this commission (CONTU), Congress amended the Act to explicitly protect computer programs as literary works.[59] Amended § 117, for

---

50. Deborah Kemp, *Limitations Upon the Software Producer's Rights*: Vault Corp. v. Quaid Software Ltd., 16 RUTGERS COMPUTER & TECH. L.J. 85, 89 (1990). The purpose of copyright law is "to promote the advancement of knowledge and learning by giving authors economic incentives (in the form of exclusive rights to their creations) to labor on creative, knowledge-enriching works." CCC Info. Serv., Inc. v. Maclean Hunter Market Reports, Inc., 44 F.3d 61, 65 (2d Cir. 1994).

51. Kemp, *supra* note 50, at 89.

52. Irvin R. Gross, *A New Framework for Software Protection: Distinguishing Between Interactive and Non-Interactive Aspects of Computer Programs*, 20 RUTGERS COMPUTER & TECH. L.J. 107, 128 (1994) (citing Act of May 31, 1790, ch. 15, 1 Stat. 124) (repealed 1802)).

53. *Id.* The Act of May 31, 1790 protected only a "book, map or chart." *Id.* at 128 n.83.

54. Copyright Act of 1909, ch. 320, § 5(a), 35 Stat. 1075, 1076 (1909) (current version at 17 U.S.C. §§ 101-1101 (1994)).

55. 17 U.S.C. § 102(a) (1994). Section 102 covers eight of the nine categories: "literary works; musical works, including any accompanying words; dramatic works, including any accompanying music; pantomimes and choreographic works; pictorial, graphic, and sculptural works; motion pictures and other audiovisual works; sound recordings; and, architectural works." *Id.* Section 103 of the Copyright Act addresses the ninth category: copyright for compilations. *See generally* 17 U.S.C. § 103 (1994).

56. H.R. REP. NO. 94-1476, at 57 (1976).

57. 17 U.S.C. § 103(b) (1994).

58. NATIONAL COMMISSION ON NEW TECHNOLOGICAL USES OF COPYRIGHTED WORKS (CONTU), 93d Cong., 2d Sess., FINAL REPORT OF THE NATIONAL COMMISSION ON NEW TECHNOLOGICAL USES OF COPYRIGHTED WORKS 1 (Comm. Print 1978) [hereinafter CONTU Report].

59. Gross, *supra* note 52, at 129 n.88 (citing Computer Software Copyright Act, Pub. L. No. 96-517, § 10(b), 94 Stat. 3028 (codified as amended at 17 U.S.C. § 117 (1994)).

example, now permits the owner of a computer program to make a backup copy of the program without copyright infringement.[60] The 1976 Act also codified case law developments under the 1909 Act by including key terms such as "selection" and "arrangement."[61] Although the CONTU Report did not expressly address "selection" and "arrangement" theories defined in the Copyright Act, it did cite three cases which followed both the "sweat of the brow"[62] and "selection and arrangement" theories.[63]

CONTU recognized that a database is an electronic version of a comprehensive compilation of facts.[64] As such, courts treat databases as "compilations" under the Copyright Act.[65] A commercially viable electronic database includes as many facts (data) as necessary to meet the needs of its enduser. A successfully compiled and designed database then allows the enduser to select and arrange the data in a fashion suitable to his or her needs.[66] A compilation's pre-existing "facts," however, remain unprotected.[67]

---

60. 17 U.S.C. § 117 (1994).

61. *See, e.g.*, Jeweler's Circular Publ'g Co. v. Keystone Publ'g Co., 281 F. 83, 88 (2d Cir. 1922).

62. A work is sufficiently copyrightable under the "sweat of the brow" or "industrious collection" theories due to the amount of labor involved in database research and design. *See* Jeweler's Circular Publ'g Co. v. Keystone Publ'g Co., 281 F. 83, 88 (2d Cir. 1922). *But see* Feist Publications, Inc. v. Rural Tele. Serv. Co., Inc., 499 U.S. 340, 353 (1991) (rejecting the sweat of the brow theory).

63. CONTU Report, *supra* note 58, at 42 n.171 (citing Leon v. Pacific Tel. & Tel. Co., 91 F.2d 484 (9th Cir. 1937); *Jeweler's*, 281 F. at 88; New York Times Co. v. Roxbury Data Interface, Inc., 434 F. Supp. 217 (D.N.J. 1977)).

64. CONTU Report, *supra* note 58, at 38. The CONTU report states that:

> [d]ictionaries, encyclopedias, and tables of numeric information are all forms of data bases [sic] which long antedate the computer, and for which copyright protection has been and will continue to be available under the copyright law. . . . This entitlement to copyright is not diminished by the fixation of the data base in a medium requiring the intervention of a computer to communicate its information content. Accordingly, a data base, whether printed in traditional hard copy or fixed in an electromagnetic medium, is protected by copyright under the terms of the new law.

Id. (footnotes omitted); *see also* Gerard J. Lewis, Jr., Comment, *Copyright Protection for Purely Factual Compilations under Feist Publications, Inc. v. Rural Telephone Service Co.: How does Feist Protect Electronic Data Bases of Facts?*, 8 SANTA CLARA COMPUTER & HIGH TECH. L.J. 169, 197 (1992).

65. *See supra* note 19 and accompanying text.

66. Lewis, *supra* note 64, at 197.

67. *See* Feist Publications, Inc. v. Rural Tele. Serv. Co., Inc., 499 U.S. 340, 344 (1991). "Section 102(b) is universally understood to prohibit any copyright in facts." *Id.* at 356. "In no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work." 17 U.S.C. § 102(b) (1994). *See also* Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 556 (1985)

Consequently, although the "compilation" of facts may be original under a sufficient showing of arrangement or selection, the individual facts comprising the compilation are not copyrightable. To establish that a compilation falls within the scope of copyright protections it must meet three criteria: (1) fixation in a tangible medium; (2) expression; and (3) originality.[68] For a database, this would include (1) fixation of the database application on diskette or CD-ROM; (2) a selection of facts representing an expression of the author; and (3) a modicum of originality exhibited via database arrangement.[69]

B. The Scope of Copyright Protection: Limitations on the Owner's Exclusive Rights

Copyright in a compilation grants certain exclusive rights to the author of the work.[70] However, in the interest of balancing protection of the owner's creative rights with public access to the work, the Copyright Act limits an owner's rights.[71] For example, libraries retain special reproduction rights in a copyrighted work.[72] Another limitation to the owner's exclusive rights is the "first sale doctrine."[73] Under this doctrine, once an author sells a copy of the work, she relinquishes control of the work and may not dictate the terms of its further disposition or transfer.[74] Hence, the "first sale" doctrine, despite the potential

---

("[n]o author may copyright his ideas or the facts he narrates").

68. 17 U.S.C. § 102(a) (1994). Section 102 provides that "[c]opyright protection subsists, . . . in original works of authorship fixed in any tangible medium of expression. . . ." *See also* Timothy D. Howell, Comment, *Intellectual Property Pirates: Congress Raises the Stakes in the Modern Battle to Protect Copyrights and Safeguard the United States Economy*, 27 ST. MARY'S L.J. 613, 631 n.49 (1996) (citing 1 PAUL GOLDSTEIN, COPYRIGHT: PRINCIPLES, LAW AND PRACTICE § 1.3, at 24-5 (1989)) ("[T]he work must be original in the sense that it was not copied from some other source; the work must consist of 'expression' and not just 'ideas'; and, the work must be embodied in a tangible medium of expression, specifically a 'copy' or 'phonorecord.'").

69. *See also* WILLIAM F. PATRY, THE FAIR USE PRIVILEGE IN COPYRIGHT LAW 517-18 (2d ed. 1995) (footnotes omitted). To constitute an original work of authorship, a database must be "(1) an independent collecting and assembling of particular preexisting material or data (2) from which a selection, coordination, or arrangement is made (3) in such a way that the resulting work as a whole 'possesses at least some minimal degree of creativity.'" *Id.*

70. Among some of the rights afforded the copyright owner in section 106 of the Copyright Act: the right to reproduce the copyright work; to prepare derivative works based upon the copyrighted work; and, distribute copies of work through "sale or other transfer of ownership, or by rental, lease, or lending." 17 U.S.C. § 106(3) (1994). See 17 U.S.C. § 106 for a full listing of the owner's exclusive rights in a copyrighted work.

71. 17 U.S.C. §§ 107-121 (1994).

72. *Id.* § 108.

73. *Id.* § 109.

74. *Id.* The first sale doctrine provides in pertinent part that "the owner of a particular copy . . . lawfully made under this title, or any person authorized by such owner, is entitled, without the authority of the copyright owner, to sell or otherwise dispose of the possession of that copy." *Id.*

risk of copyright infringement, allows an enduser to lend computer software to friends for evaluative purposes or to sell the software to a third party.

One of the most extensive limitations on the use of a copyrighted work is found in § 107: fair use.[75] A defendant may assert that his use of the copyrighted work was fair by applying the four factors defined in § 107:[76]

> In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include—(1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work.[77]

The preamble, however, to § 107[78] indicates that these fair use factors are not exhaustive.[79] Courts have applied other factors to determine fair use which Congress condoned in its 1992 amendment of § 107.[80] As such, courts have applied additional factors such as "defendant's good faith or lack thereof" or "wrongful denial of exploitative conduct towards the work of another" to determine fair use of a copyrighted work.[81] Yet, when a court applies the fair use doctrine to compilations, it may only assess the amount and substantiality of the *protectable* portions used; thus, after *Feist*, this leaves only the "selection or arrangement" of the work—the unprotectable content (facts) may not be considered.[82]

Finally, § 301 of the Copyright Act[83] calls for preemption of a state law that grants "rights that are equivalent to any of the exclusive rights within the general scope of copyright as specified by § 106 in works of authorship that . . . come

---

75. 17 U.S.C. § 107. Fair use is a codification of judicial doctrine. *See* PATRY, *supra* note 69, at 413. "Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies . . . , for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright." 17 U.S.C. § 107.

76. *Id.*

77. *Id.*

78. *Id.*

79. PATRY, *supra* note 69, at 568. "The preamble to Section 107 uses the phrase 'including such,' the term 'such as,' and the word 'include' . . . . The listing of fair use purposes and factors is thus 'not intended to be exhaustive.'" *Id.* (quoting Harper & Row Publishers, Inc. v. Nation Enters., 471 U.S. 539 (1985)).

80. *Id.* (quoting H.R. REP. NO. 836, at 9-10 (1992)).

81. *Id.* at 568-89.

82. *Id.* at 518. This secures the protection of originality while leaving the facts contained in the compilation in the public domain. In a database, the court may assess the fair use of the arrangement but not of the facts themselves. *Id.*

83. 17 U.S.C. § 301 (1994).

within the subject matter of copyright as specified by §§ 102 and 103 . . . ."[84] Generally, copyright preemption requires: (1) the work fall within the subject matter of copyright as defined in §§ 102 and 103; and (2) the state right granted must be "equivalent" to one or more rights granted by § 106 of the Copyright Act.[85] Consequently, a state law protecting an electronic databases' content should be preempted pursuant to § 301.[86] However, if a state law requires an "extra element" not required by copyright, then the right does not lie "within the general scope of copyright" and escapes § 301 preemption.[87] The extra element test is met if the state cause of action is "qualitatively different from an action for copyright."[88]

C. Advancing the Goals of Copyright

By enacting the copyright provisions, Congress met the goals of the Constitutions's Copyright Clause. The fundamental policy behind the Copyright Clause is best represented in *Mazer v. Stein*: "The economic philosophy behind the clause empowering Congress to grant patents and copyrights is the conviction that encouragement of individual effort by personal gain is the best way to advance public welfare through the talents of authors and inventors in 'Science and useful Arts.'"[89] The Copyright Clause also serves to reward the public.[90] After a period of time, the Copyright Act vests the public with rights to the work unencumbered by copyright infringement laws.[91]

---

84. 17 U.S.C. § 301(a) (1994).

85. National Car Rental Sys., Inc. v. Computer Assoc. Int'l, Inc., 991 F.2d 426, 428 (8th Cir. 1993) (quoting Harper & Row, Publishers, Inc. v. Nation Enters., 723 F.2d 195, 200 (2d Cir. 1983). *See also* III PAUL GOLDSTEIN, COPYRIGHT §15.12.1 (2d ed. 1996) (adding a third element: fixation in a tangible medium).

86. GOLDSTEIN, *supra* note 83, § 15.12.1 n.4.

87. 1 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT, § 1-01[B][1], at 1-15 (2d ed. 1996) [hereinafter NIMMER]. Nimmer asserts that if other elements "are required instead of, or in addition to, the acts of reproduction, performance, distribution or display," then the state created right may not be preempted. *See also* Harper & Row, Publishers, Inc. v. Nation Enters., 501 F. Supp. 848 (S.D.N.Y. 1980), *aff'd*, 723 F.2d 195 (2d Cir. 1983), *rev'd on other grounds*, 471 U.S. 539 (1985) (Second Circuit court found that the state law was not preempted because the equivalent rights test was not satisfied given the extra elements found in the conversion and interference with contractual relations claims).

88. *National Car Rental*, 991 F.2d at 431.

89. 347 U.S. 201, 219 (1954).

90. Kemp, *supra* note 52, at 89. "The idea is that the public good is served by protecting creators in their intellectual property." *Id.* at 89 n.15.

91. 17 U.S.C. § 302(a) (1994). Section 302 generally provides that the copyright protection "endures for a term consisting of the life of the author and fifty years after the author's death." *Id.* For works created within the scope of employment (a "work made for hire"), the public acquires rights to the work after a period of seventy-five years from date of publication or 100 years after the work is created, "whichever expires first." 17 U.S.C. § 302(c) (1994).

"The primary objective of copyright is not to reward the *labor* of authors, but '[t]o promote the Progress of Science and useful Arts.'"[92] To this end, the Copyright Act reserves "facts" for the public domain, allowing authors to create and build upon them so as to enrich the public. "[C]opyright does not prevent subsequent users from copying from a prior author's work those constituent elements that are not original—for example . . . facts, or materials in the public domain—as long as such use does not unfairly appropriate the author's original contributions."[93] The Copyright Act serves to protect this building block approach "without impeding the public's access to that information which gives meaning to our society's highly valued freedom of expression."[94]

These policies apply to compilations and computer databases. "Maximization of public access to information contained in automated data bases [sic] is . . . a significant goal of [the] national information policy."[95] Further, encouraging database development for dissemination to the public supports the underlying policies of the Copyright Act.[96] However, "[t]here is an inherent conflict in the competing interests of the author and the public. Protecting the rights of the author could be detrimental to the public's interest in access to the work of authorship."[97] The Copyright Act "seeks to balance these potentially conflicting interests."[98] The delicate balance of encouraging database development for the benefit of the public has in some respects become skewed in favor of public access.[99] Database development comes with significant costs

---

92. Feist Publications, Inc. v. Rural Tele. Serv. Co., Inc., 499 U.S. 340, 349 (1991) (quoting U.S. CONST. art. I, § 8, cl. 8) (emphasis added). The Court reasoned that this underlying policy assures the author protection of his expression but also encourages other authors to "build freely upon the ideas and expressions conveyed by a work." *Id.* at 350.

93. Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 548 (1985). The court in *Vault v. Quaid* directs this underlying policy toward computer technology: "After acknowledging the importance of balancing the interest of proprietors in obtaining 'reasonable protection' against the risks of 'unduly burdening users of programs and the general public' . . . [CONTU] recommended the repeal of section 117. . . and the enactment of a new section 117 which would proscribe the unauthorized copying of computer programs but permit a 'rightful possessor' of a program to make or authorize the making of another copy . . . ." Vault Corp. v. Quaid Software Ltd., 847 F.2d 255, 259-60 (5th Cir. 1988) (footnotes omitted).

94. Harper & Row, Publishers, Inc. v. Nation Enters., 723 F.2d 195, 202 (2d Cir. 1983).

95. CONTU Report, *supra* note 58, at 38 n.159 (citing REPORT TO THE PRESIDENT OF THE UNITED STATES ON NATIONAL INFORMATION POLICY 70 (1976)).

96. CONTU Report, *supra* note 58, at 38. The report encourages the registration of the compilation work so as to make the public aware of these works and gain access to its useful information.

97. Kemp, *supra* note 52, at 89.

98. *Id.*

99. The lower court in *ProCD* left the database unprotected: "the Select Phone™ listings themselves are uncopyrightable." ProCD, Inc. v. Zeidenberg, 908 F. Supp. 640, 647 (W.D. Wis. 1996), *overruled by* ProCD, Inc. v. Zeidenberg, 86 F.3d 1447 (7th Cir. 1996) (also finding database uncopyrightable pursuant to *Feist*.).

both to the developer and the public.[100] Without database protection, the developer might be discouraged from compiling comprehensive repositories of information—to the detriment of the public. The reduction in the available repositories would negatively impact the "Progress of Science and useful Arts."[101] Therefore, the balance should be restored in line with the policies set forth by the courts, Congress and the Constitution.

III. THE STATE OF DATABASE PROTECTION UNDER COPYRIGHT LAW

Copyright protection for databases has weakened. The Supreme Court's decision in *Feist* eroded database protection when it refuted the "sweat of the brow" theory.[102] Even though the Seventh Circuit in *ProCD* ultimately protected the database, the court did so outside the confines of copyright law.[103] Database developers have relied upon state contract law to protect their valuable investment and, in the Seventh Circuit at least, this reliance was well placed.

Ironically, the same result could have been reached in *Feist* without eliminating the "sweat of the brow" theory. The Court could have applied the fair use doctrine—a long standing judicially-created device to protect both creators and users of copyrighted works.[104] Feist Publications copied a nominal amount of data from the Rural directory to supplement Feist's area-wide directory.[105] This would constitute a fair use because "[t]he only material Feist took from Rural's white pages were raw facts which were themselves uncopyrightable . . . . [M]erely using [these raw facts] as a reference work could easily suffice as fair use under § 107 of the Copyright Act."[106]

After *Feist*, however, only the selection and arrangement theory remain to protect an electronic database.[107] A database designer must creatively select and arrange the data to secure copyright protection in this investment. In light of *Feist*, the *ProCD* court, in passing, noted that the SelectPhone™ database was not copyrightable.[108] However, a more extensive analysis might have revealed

---

100. ProCD argued that "it is unfair and commercially destructive to allow [Zeidenberg] to take the information [ProCD] assembled with a significant investment of time, effort and money and use it for commercial purposes without paying any compensation. . . ." *ProCD*, 908 F. Supp. at 646.

101. U.S. CONST. art. I, sec. 8, cl. 8. *See* Paul J. Heald, *The Vices of Originality*, 1991 SUP. CT. REV. 143, 152 (1991).

102. Feist Publications, Inc. v. Rural Tele. Serv. Co., 499 U.S. 340, 349 (1991).

103. *ProCD*, 86 F.3d at 1455.

104. *See supra* note 75 and accompanying text.

105. *Feist*, 499 U.S. at 343. The Court noted that Feist only selected 1309 listings (four of which were dummy listings) of Rural's 7700 listings. *Id.* at 343-44. This constituted seventeen percent of Rural's factual listing.

106. Heald, *supra* note 99, at 147 (footnotes omitted).

107. *See Feist*, 499 U.S. at 362 and *supra* note 25 and accompanying text.

108. *ProCD*, 86 F.3d at 1449. Presumably, the court did not pursue this analysis as the *Feist* decision found that white pages were uncopyrightable because they lacked the requisite originality

that SelectPhone™ *could* have been protected under the "selection" and "arrangement" criteria which *Feist* upheld.[109] The court alludes to such protection: "[w]e may assume that this database cannot be copyrighted, *although* it is *more complex,* contains more information (nine-digit zip codes and census industrial codes), is *organized differently*, and therefore is *more original* than the single alphabetical directory at issue in [*Feist*]."[110] The selection and arrangement of the ProCD database met the "modicum of creativity" required to find originality in the work.[111] Assuming originality in the database, the court might have found that Zeidenberg copied a substantial amount of the database *arrangement*, thereby infringing ProCD's copyright.[112] However, instead of copyright, the *ProCD* court protected the database through a shrinkwrap license—contract law.[113]

A competitive database requires the developer to provide an all-inclusive repository of information. In turn, the enduser "selects" only that information pertinent to his research. A comprehensive database therefore, penalizes the developer because the *selection* element is not met due to the all-inclusiveness of the data. Although the *arrangement* of the database qualifies the work for copyright protection, extensive protection of the entire database—the facts—is lacking. Only the arrangement is protected and not the data comprising the database. A competitor can then take the data by simply "rearranging" the facts in a new compilation.[114] Yet, this is the very thing the database vendor seeks to protect—the electronic compilation of facts; therein resides the company's exhaustive and expensive research.

To protect electronic databases from misappropriation, the facts need protection.[115] However, "[t]he most fundamental axiom of copyright law is that

---

required by the Constitution. *Feist*, 499 U.S. at 362-63. The ProCD database represents an electronic version of the white pages.

109. *Feist*, 499 U.S. at 362. Some commentators have asserted that electronic databases would still benefit from copyright protection as the arrangement of the database would sufficiently satisfy the originality requirement for copyrighted compilations. *See, e.g.*, Litman, *supra* note 30, at 609 ("Unlike white pages directories, computer databases are surely sufficiently original to merit copyright protection"); Lewis, *supra* note 64, at 205 ("The selection theory is well-suited for electronic data bases because it reflects the intrinsic value of databases and the selection process used to create and update data bases."). *But see* Hayden, *supra* note 30, at 219 (author asserts the *Feist* standard of selection and arrangement is difficult to apply to computer databases given their comprehensiveness in data selection).

110. *ProCD*, 86 F.3d at 1449 (emphasis added).

111. *Feist*, 499 U.S. at 346.

112. Zeidenberg copied the entire SelectPhone™ database to the Internet for his customers to access. *ProCD*, 86 F.3d at 1450. This would constitute the copying of the arrangement, hence, the creativity of ProCD developers.

113. *ProCD*, 86 F.3d at 1449.

114. Litman, *supra* note 30, at 609. "[A] competitor would be infringing no copyright if it simply stole the data and left the base." *Id.* (footnote omitted).

115. Professor Ginsburg characterizes comprehensive databases as "low authorship works."

'[n]o author may copyright his ideas or the facts he narrates.'"[116] However, the lack of protection for these facts within an electronic database may lead to low incentives for creating beneficial works.[117] In the end, the public suffers. Therefore, the goal of copyright reform for compilations should center around "maximiz[ing] public welfare through a coherent system of incentives."[118]

The goal of the Copyright Act and the Copyright Clause seeks to "promote" the creation of works. *Feist*, by eroding database protection, discourages the creation of such useful tools.[119] The technology industry has therefore attempted to strengthen database protection through other avenues—the shrinkwrap license. Fortunately, for ProCD, a sympathetic court enforced ProCD's shrinkwrap license resulting in an exclusive right to the database. Such exclusive rights, while protecting the developer (creator), interfere with the rights of endusers. Shrinkwrap licenses and other methods of database protection have emerged as a result of eroding copyright protection for databases. However, alternative methods must continue to honor the underlying policy of the Constitution: the promotion of "Science and useful Arts."

IV. AN ANALYSIS OF VARIOUS EFFORTS TO REFORM DATABASE PROTECTION

A. Shrinkwrap Licenses and the Copyright Act

Prior to the 1980 Copyright amendment, copyright protection for software programs was as uncertain as database protection is today. To advance an

---

Low authorship works are "personality-deprived information compilations such as directories, indexes and data bases [sic]." Jane C. Ginsburg, *Creation and Commercial Value: Copyright Protection of Works of Information*, 90 COLUM. L. REV. 1865, 1866 (1990).

116. Feist, 499 U.S. at 344 (citing Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 556 (1985).

117. *See generally* Robert C. Denicola, *Copyright in Collections of Facts: A Theory for the Protection of Nonfiction Literary Works*, 81 COLUM. L. REV. 516 (1981) (contending that low protection leads to low incentives which would deplete the level of works available to the public and therefore advocates rewarding industrious efforts in compiling low authorship works); Ginsburg, *supra* note 113 (concluding that the lack of protection of low authorship works would lead to low incentive to create such works and therefore less variety of low authorship works). *But see* Litman, *supra* note 30, at 646 (Professor Litman rejects statutory provisions which protect facts within compilations. She argues that the only "legitimate justification for conferring statutory protection on facts is ensuring that public access to factual compilations is enhanced."); Raymond T. Nimmer, *Information Age in Law: New Frontiers in Property and Contract*, 68 N.Y. ST. B. J. 28 (1996) (the author recognizes that "[a] balance must exist between private control . . . and public access or use"); L. Ray Patterson & Craig Joyce, *Monopolizing the Law: The Scope of Copyright Protection for Law Reports and Statutory Compilations*, 36 UCLA L. REV. 719 (1989) (contending that low authorship works are overprotected allowing public domain materials to be monopolized by authors who compile facts).

118. Heald, *supra* note 99, at 154.

119. *See generally* Ginsburg, *supra* note 113; Heald, *supra* note 99.

appearance of protection, the software industry developed the "shrinkwrap license."[120] Under a shrinkwrap license, for example, an enduser agrees to use the software on his system only and to refrain from copying the software.[121] Such licenses projected the image of copyright protection in the electronic work.[122]

1. Case Law Interpretations of Shrinkwrap Licenses.—Recognizing the limitations of copyright law, ProCD protected its database through the use of a shrinkwrap license.[123] ProCD included an electronic shrinkwrap license to which Zeidenberg assented before using the database application.[124] While the district court in *ProCD* followed judicial precedent in holding the shrinkwrap license unenforceable,[125] the Seventh Circuit reversed this holding and further held that the shrinkwrap license was not preempted by the federal Copyright Act.[126]

Only three lower courts have addressed the enforceability of shrinkwrap licenses.[127] In a battle of the forms case, the Third Circuit held that a shrinkwrap license may not modify the terms of a previous contract. Thus, as a matter of contract law, the shrinkwrap license was unenforceable.[128] Similarly, in *Arizona Retail Systems, Inc. v. The Software Link, Inc.*, a district court held the shrinkwrap license unenforceable because its terms included "additional terms" which "materially altered" the contract without the express acceptance of the party.[129] In 1988, the Fifth Circuit concluded that a shrinkwrap license, although

---

120. These licenses are sometimes referred to as "box top" licenses but serve the same purpose as shrinkwrap licenses.

121. Some commentators classify shrinkwrap licenses as "contracts of adhesion" because they "deprive the purchaser of important rights which he has not relinquished voluntarily." Page M. Kaufman, Note, *The Enforceability of State "Shrink-Wrap" License Statutes in Light of* Vault Corp. v. Quaid Software, Ltd., 74 CORNELL L. REV. 222, 235 (1988).

122. Today, an electronic version of the shrinkwrap license has evolved where the vendor creates a separate "licensing" screen which informs the user of the license to which the user much assent before using the product. *See, e.g.*, *ProCD*, 86 F.3d at 1450.

123. ProCD, Inc. v. Zeidenberg, 908 F. Supp. 640, 650 (W.D. Wis. 1996). "Before marketing SelectPhone™, plaintiff recognized the potential limitations of copyright law and included an agreement in the software package that sought to impose limitations concerning the distribution and use of the telephone listings." *Id.*

124. This electronic license was duplicated in detail in the user manual. *ProCD*, 86 F.3d at 1450.

125. *ProCD*, 908 F. Supp. at 644. The district court held that the shrinkwrap license was unenforceable and preempted by copyright law. *Id.*

126. *ProCD*, 86 F.3d at 1455.

127. *See* Step-Saver Data Sys., Inc. v. Wyse Tech., 939 F.2d 91 (3d Cir. 1991); Vault Corp. v. Quaid Software, Ltd., 847 F.2d 255 (5th Cir. 1988); Arizona Retail Sys., Inc. v. The Software Link, Inc., 831 F. Supp. 759 (D. Ariz. 1993).

128. *Step-Saver*, 939 F.2d at 105. The court applied U.C.C. § 2-207 to determine that disclaimers of warranty and liability (contained in the license) between a software developer and a value-added retailer were additions to the contract which materially altered the agreement and therefore could not be incorporated in the contract. *Id.*

129. *Arizona Retail*, 831 F. Supp. at 764-66. This case opened the shrinkwrap license door

compliant with the state's Software License Act, was preempted by federal copyright law and therefore unenforceable.[130]

Notwithstanding prior decisions, the Seventh Circuit broke new ground by enforcing the restrictive license in *ProCD*.[131] Judge Easterbrook afforded ProCD rights it did not have under copyright law.[132] Under *Feist*, the court concluded

---

slightly by acknowledging the validity of shrinkwrap licenses when accompanied by a software evaluation diskette. The court held that if Arizona Retail "requested an evaluation diskette and then, by keeping the live disk, agreed to purchase the copy of [the software] that accompanied the evaluation diskette after evaluating [the software], the license agreement applies to the initial transaction. Under such facts, the contract was . . . formed . . . after [Arizona Retail] opened the shrink wrap [sic] on the live version of [the software] which [Arizona Retail] had notice would result in a contract being formed." *Id.* at 763.

130. *Vault Corp.*, 847 F.2d at 270. The Software License Act limited the user's right to reverse engineer the software. *Id.* at 257 n.2. *See infra* note 157 and accompanying text.

131. *ProCD*, 86 F.3d at 1450. The license in *ProCD* limited the use of the database application to "non-commercial" purposes. *Id.* The license, which appeared on the screen when the enduser accessed the software program, contained the following information:

> The listings on this product are licensed for authorized users only. The user agreement provides that *copying of* the software and *the data* may be done only for individual or personal use and that distribution, sublicense or lease of the software or the data is prohibited. . . .[Y]ou will not make the Software or the Listings in whole or in part available to any other user in any networked or time-shared environment, or transfer the Listings in whole or in part to any computer other than the computer used to access the Listings.

ProCD, Inc. v. Zeidenberg, 908 F. Supp. 640, 645 (W.D. Wis. 1996) (emphasis added).

132. "[W]hile contractual provisions can be helpful in determining the scope of a creator's proprietary interest in software, a 'license' of computer software that is nothing more than a disguised sale should not be enforceable as a 'back door' way of expanding the Copyright Act's protection." Thomas Lee Hazen, *Contract Principles as a Guide for Protecting Intellectual Property Rights in Computer Software: The Limits of Copyright Protection, the Evolving Concept of Derivative Works, and the Proper Limits of Licensing Arrangements*, 20 U.C. DAVIS L. REV. 105, 112 (1986). *See also* Kaufman, *supra* note 119, at 224 (asserting that shrinkwrap licenses should remain unenforceable as they "purport to grant rights that are either equivalent to or in direct conflict with rights granted under the Copyright Act"). Another commentator suggests that shrinkwrap licenses modify copyright law in several ways (evaluating the *Vault v. Quaid* licensing agreement):

> [1] [T]he license prevents the user from selling or otherwise disposing of that particular copy of the software. This conflicts with the "first sale" doctrine in copyright law, which (with some statutory restrictions on renting and lending in the case of computer software) allows purchasers of a particular copy of a work to give or sell that copy to another. [2] [T]he license prevents the user from copying, modifying, translating, or converting the program for any purpose. This provision at least potentially conflicts with section 117 of the Copyright Act . . . . [3] [T]he license provides that the user may not "decompile or disassemble" the program for any purpose. This provision conflicts with the majority rule in copyright law that users may "reverse engineer" computer

that ProCD did not have copyright protection in the database.[133] Nonetheless, the enforcement of the shrinkwrap license afforded ProCD exclusive rights in the database. The court effectively (1) extended database protection by preventing the enduser from copying the database application; (2) by-passed copyright's first sale doctrine; (3) permitted an ambiguous acceptance of contract terms following the purchase of a product; and, (4) ignored the doctrine of fair use.[134]

Unlike copyright law, the shrinkwrap license prohibited the user from copying the data contained in the database or from leasing the application.[135] The ProCD license prohibited the lease or other disposition of the database application to another party which is in direct conflict with the first sale doctrine of copyright law.[136] Prohibiting copying of the database application may interfere with the Copyright Act's "fair use" doctrine, especially if the database were used for educational purposes.[137] The ProCD shrinkwrap license secured protection in the raw data whereas copyright law does not protect facts contained in a compilation.[138] Hence, although *ProCD* successfully protected the database, it did so at the expense of the enduser's rights—rights which copyright law protects.

Further, the Seventh Circuit, by enforcing the shrinkwrap license, endorsed the use of ambiguous terms for contract acceptance after the purchase of a product.[139] Relying on section 2-204(1) of the Uniform Commercial Code, the court found that ProCD, as "master of the offer," displayed its license terms on the screen and "invited" acceptance by requiring the buyer to perform a certain act—pressing "Enter"—before proceeding.[140] Although the buyer had notice of

---

programs when necessary to have access to unprotected ideas contained in those programs.

Mark A. Lemley, *Intellectual Property and Shrinkwrap Licenses*, 68 S. CAL. L. REV. 1239, 1246-47 (1995).

133. ProCD, Inc. v. Zeidenberg, 86 F.3d 1447, 1449 (7th Cir. 1996).

134. *See generally id.* at 1447.

135. ProCD, Inc. v. Zeidenberg, 908 F. Supp. 640, 645 (W.D. Wis. 1996). See *supra* note 132 for the pertinent text of the ProCD license agreement.

136. *See* 17 U.S.C. § 109(a) (1994). "The rightful owner of a copy of a copyrighted work has the right to sell that copy, but nothing more." Hazen, *supra* note 130, at 112. Professor Hazen further asserts that the computer industry relies on licensing agreements to avoid the first sale doctrine. *Id.*

137. *See* 17 U.S.C. § 107 (1994).

138. Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 556 (1985).

139. Some courts have held that a shrinkwrap license is unenforceable as a matter of contract law because the user has not assented to the terms prior to the purchase of the product. *See, e.g.*, Step-Saver Data Sys., Inc. v. Wyse Tech., Inc., 939 F.2d 91 (3d Cir. 1991); Arizona Retail Sys. v. Software Link, Inc., 831 F. Supp. 759 (D. Ariz. 1993).

140. ProCD, Inc. v. Zeidenberg, 86 F.3d 1447, 1452 (7th Cir. 1996). "[U.C.C.] § 2-204(1) [states that] '[a] contract for sale of goods may be made in any manner sufficient to show agreement, including conduct by both parties which recognizes the existence of such a contract' . . . ProCD proposed a contract that a buyer would accept by using the software after having an

a restrictive license because the packaging displayed such a warning, the explicit terms of the license were unknown to the buyer until he purchased the product.[141] However, the court reasoned that "ProCD extended an opportunity to reject [the product] if a buyer [found] the license terms unsatisfactory."[142] Ultimately, Judge Easterbrook construed the U.C.C. in such a manner so as to grant protection in the ProCD database.

2. The U.C.C. Draft and Its Interpretation of Shrinkwrap Licenses.—The American Law Institute (ALI) along with the National Conference of Commissioners on Uniform State Laws Committee supports Judge Easterbrook's rationale in its new draft section of the Uniform Commercial Code ("U.C.C."), section 2B-308, entitled "Mass Market Licenses."[143] Section 2B-308 of the draft validates shrinkwrap licenses.[144] Pursuant to section 2B-308, if the buyer assents to the terms in a manner prescribed by the licensor, the shrinkwrap license is enforceable.[145] In *ProCD*, simply selecting "OK" from the screen displaying key

---

opportunity to read the license at leisure. This Zeidenberg *did.* He had no choice, because the software splashed the license on the screen and would not let him proceed without indicating acceptance." *Id.* (emphasis added).

141. *Id.* at 1450. It is this process which commentators find disturbing. One commentator identifies two major problems with enforcing shrinkwrap licenses: "[1] that consideration has passed before the license terms are placed before the user, and [2] there is an ambiguous act of acceptance by the user to the license terms." Joel Rothstein Wolfson, *Information Transactions on the Information Superhighway: It's Not Just Software Law Anymore*, 6 NO. 11 J. PROPRIETARY RTS. 2 (November 1994).

142. *ProCD*, 86 F.3d at 1452-53. The court applied section 2-606 of the U.C.C., which defines "acceptance of goods" as support for its reasoning. Any buyer who is dissatisfied with the terms of the agreement "can prevent formation of the contract by returning the package, as can any consumer who concludes that the terms of the license make the software worth less than the purchase price." *Id.* at 1452. *See also* Hill v. Gateway 200, Inc., 105 F.3d 1147, 1148-49 (7th Cir. 1997) (Easterbrook maintains that parties agree to all the terms of a contract "in a box" if a party has an opportunity to accept or return the product).

143. Henry Beck, *Uniform Commercial Code Article 2B—Licenses*, 478 PLI/PAT. 103, 174 (1997) [hereinafter U.C.C. Draft]. This article contains the body of the U.C.C. Draft Article 2B with Reporter's notes. A copy of the applicable "Mass Market Licenses" section (§ 2B-308), as of the publication of this Note, is replicated in the Appendix. The U.C.C. Draft defines a mass market license as follows:

> a standard prepared for and used in a retail market for information which is directed to the general public as a whole under substantially the same terms for the same information, if the licensee is an end user and acquired the information in a transaction under terms and in a quantity consistent with an ordinary transaction in the general retail distribution. The term includes consumer contracts.

Id. 2B-102(a)(25).

144. *Id.*

145. *Id.* at 174. "[A] party adopts the terms of a mass market license if, the party agrees or manifests assent to the mass market license before or in connection with the initial use of or access to the information." *Id.* (citing § 2B-308(a)).

license terms would have been sufficient to manifest assent under section 2B-308 rendering the shrinkwrap license enforceable.

The Drafting Committee characterizes the shrinkwrap license as a standard from contract and, as such, the shrinkwrap license must comply with standard procedures outlined in section 2B-307.[146] Thus, the licensee manifests assent through his signature or conduct. The Committee defines "manifest assent" as an opportunity to review the license agreement before assenting to the terms of the license.[147] The electronic license agreement in *ProCD*, for example, presented the enduser with an opportunity to review the pertinent terms of the agreement prior to acceptance.[148] This is not the only safeguard available to the purchaser of a product with a shrinkwrap license. If a term in the license causes undue surprise or deviates from industry standards, that term is unenforceable regardless of consent.[149]

---

146. *Id.* at 172 (citing § 2B-307(a)). As of the publication of this Note, section 2B-307 reads in its entirety as follows:

> (a) Except as otherwise provided in subsection (c) and Sections 2B-308 and 2B-309, a party adopts the terms of a record, including a standard form, if the party agrees to or manifests assent to the record before or in connection with the initial use of or access to the information. If agreement or assent to a record does not occur by that time, but the parties commence performance or use the information with the expectation that their agreement will be later represented in whole or in part by a record that the party has not yet had an opportunity to review or that has not yet been completed, a party adopts the terms of the later record if the party agrees to or manifests assent to that record.
>
> (b) A term adopted under subsection (a) becomes part of the contract without regard to the knowledge or understanding of the individual term by the party assenting to the record whether or not the part read the record.
>
> (c) A term of a record which is unenforceable for failure to satisfy a requirement of another provision of this article, such as a provision that expressly requires use of conspicuous language or manifested assent to the term, is not part of the contract.

147. *Id.* at 174-75.

> This concept adopts procedural safeguards allowing the party bound by the standard form an opportunity to review terms and to reject the contract if the terms are not acceptable. The two safeguards are in the concept of "opportunity to review" (see 2B-114) and "manifests assent" (see 2B-113). Under these definitions, a party cannot manifest assent to a form or a provision of a form unless it has had an opportunity to review that form before being asked to react.

Id.

148. ProCD v. Zeidenberg, 86 F.3d 1447, 1452 (7th Cir. 1996).

149. *See* U.C.C. Draft, *supra* note 141, at 174 (citing § 2B-308(b)(1)). Section 2B-308(b) provides in relevant part:

> [A] term does not become part of the contract if the term creates an obligation or imposes a limitation which: (1) the party proposing the form should know would cause an ordinary and reasonable person acquiring this type of information and receiving the

Although assenting to terms in a contract before review is a significant issue presented by shrinkwrap licenses, the most disturbing issue is its direct interference and conflict with federal copyright law. Opponents of this provision are concerned with diminished user's rights resulting from shrinkwrap license enforcement.[150] They support the addition of a provision that provides for the abatement of any terms in the contract that are inconsistent with copyright law.[151] Others maintain that such incorporation of copyright law into contract law is ill-advised.[152] These advocates mention that Copyright's preemption clause and fair use doctrine equip the judiciary with adequate tools that can address these concerns.[153] However, the preemption clause and fair use doctrine were scantily applied in *ProCD* affording protection that would not otherwise be afforded a database under copyright law.

Shrinkwrap licenses are typically designed to limit the user's rights to the database application which copyright law would permit—for example, limiting the duplication of data within the database. Consequently, the Committee has offered a form of database protection outside the realm of copyright law. In *ProCD*, the shrinkwrap license protected the database from commercial copying which is essentially protection against unfair competition.[154] Database developers are concerned with a commercial user copying their work and developing a competing product at a lower cost. This is what shrinkwrap licenses ultimately seek to prevent. However, the U.C.C. Draft jeopardizes the

---

> form to refuse the license if that party knew that the license contained the particular form . . . ."

Id.

150. *Copyrights: ALI Votes That New UCC Provision Must Be Consistent with Copyright Act*, BNA Patent, Trademark & Copyright Law Daily, June 6, 1997, at D2 (opponents are concerned with the ability of this provision to contract copyright rights away as well as the loss of rights the education and library sectors enjoy under copyright law).

151. *Copyright: ALI Wants Law on Information Licensing to Include Current Copyright Statute*, BNA Patent, Trademark & Copyright Law Daily, May 22, 1997, at D3. The proposed provision provides that a "term that is inconsistent with 17 U.S.C. Section 102(b) or with the limitations on exclusive rights contained in 17 U.S.C. Section 107-112 and 117 cannot become part of a contract under this [Mass Marketing license] section." *Id.*

152. *Id.*

153. *Id.*

154. Understanding the ease at which electronic works can be copied and misappropriated, CONTU assessed the value of unfair competition laws.

> The common law doctrine of unfair competition of the misappropriation variety is based upon the principle that one may not appropriate a competitor's skill, expenditure, and labor. . . . While there is a small body of federal unfair competition law, it is largely a state doctrine with the same lack of national uniformity that besets trade secrecy. Although unfair competition may provide relief ancillary to copyright in certain situations, its scope is not as broad, and it seems unlikely that it alone could provide sufficient protection against the misappropriation of programs.

CONTU Report, *supra* note 58, at 18.

delicate balance between protecting an author's right (misappropriation) and the public's right to access information in a work. Can traditional judicial doctrines of fair use or copyright preemption retain the balance sought by the founders of the Constitution?

3. An Analysis of Preemption and Shrinkwrap Licenses.—Although shrinkwrap licenses may provide database protection under contract law, § 301 of the Copyright Act preempts such licenses.[155] Many law review articles support the unenforceability of shrinkwrap licenses for precisely this reason.[156] While some courts wrestle with the contractual implications of shrinkwrap licenses,[157] others carefully construe the shrinkwrap license under federal preemption.[158] *Vault Corp. v. Quaid Software Ltd.*[159] is the most notable case addressing the preemption issue and shrinkwrap licenses.

Vault created a copy protection program ("Prolock") designed to prevent the unlawful copying of software by "locking" the protected software.[160] In other words, a software developer could use Prolock to prevent his software from being copied. Vault sold Prolock under a shrinkwrap license which prohibited the reverse engineering of the software.[161] Nonetheless, Quaid purchased the software and reverse engineered the package to unlock the "locking" function. Quaid's product, which disabled Prolock, was sold as "Ramkey."[162]

---

155. *See* 17 U.S.C § 301 (1994).

156. *See, e.g.*, David A. Einhorn, *Box-Top Licenses and the Battle-of-the-Forms*, 5 SOFTWARE L.J. 401 (1992); Gary W. Hamilton and Jeffrey C. Hood, *The Shrink-Wrap License—Is It Really Necessary?*, 10 No. 8 COMPUTER LAW. 16 (1993); Hazen, *supra* note 130; David A. Rice, *Public Goods, Private Contract and Public Policy: Federal Preemption of Software License Prohibitions Against Reverse Engineering*, 53 U. PITT. L. REV. 543 (1992); Karen Puhala, Note, *The Protection of Computer Software Through Shrink-Wrap License Agreements*, 42 WASH. & LEE L. REV. 1347 (1985). *But see* Robert W. Gomulkiewicz and Mary L. Williamson, *Article: A Brief Defense of Mass Market Software License Agreements*, 22 RUTGERS COMPUTER & TECH. L.J. 335 (1996); Robert P. Merges, *Intellectual Property and the Costs of Commercial Exchange: A Review Essay*, 93 MICH. L. REV. 1570 (1995).

157. *See, e.g.*, ProCD v. Zeidenberg, 86 F.3d 1447 (7th Cir. 1996); Arizona Retail Sys., Inc. v. The Software Link, Inc., 831 F. Supp. 759 (D. Ariz. 1993).

158. Vault Corp. v. Quaid Software, Ltd., 847 F.2d 255 (5th Cir. 1988).

159. *Id.*

160. *Id.* at 256.

161. *Id.* at 256-57. The process of reverse engineering is complicated. Computer programs consist of source code (COBOL, FORTRAN . . .) which programmers are trained to understand. The source code is translated into "object code," or machine-readable language. "Reverse engineering, as the name implies, involves going backwards from a finished product and determining how the program works. One method of reverse engineering involves the use of disassemblers, which allow someone to convert object code into a form more easily read by humans [source code]." Sega Enter. Ltd. v. Accolade, Inc., 785 F. Supp. 1392, 1394-95 (N.D. Cal. 1992), *aff'd in part, rev'd in part*, 977 F.2d 1510 (9th Cir. 1992).

162. *Vault*, 847 F.2d at 257.

Notwithstanding its compliance with the Louisiana Software License Act,[163] the Fifth Circuit held the shrinkwrap license unenforceable because federal law preempted this state law.[164] The court's conclusion that a state could not expand the rights of authorship granted by the copyright law has been followed in subsequent cases.[165]

Interestingly, *Vault* did not apply § 301 of the Copyright Act but relied instead upon Supremacy Clause preemption.[166] Section 301 of the Copyright Act calls for preemption if (1) the work falls within the subject matter of copyright as defined in §§ 102 and 103 of the Copyright Act; and (2) the state right granted is "equivalent" to one or more rights granted by § 106 of the Copyright Act.[167] Section 301 expressly declares Congress' intent: to preempt state law where "equivalent" rights to copyright are extended.[168] The court in *Vault*, however,

---

163. *Id.* at 268-70 (citing LA. REV. STAT. ANN. §§ 51:1961-1965 (West 1987)).

164. *Id.* at 270.

> The provision in Louisiana's License Act, which permits a software producer to prohibit the adaptation of its licensed computer program by decompilation or disassembly, conflicts with the rights of computer program owners under § 117 and clearly 'touches upon an area' of federal copyright law. For this reason, and the reasons set forth by the district court, we hold that at least this provision of the Louisiana's License Act is preempted by federal law, and thus that the restriction in Vault's license agreement against decompilation or disassembly is unenforceable.

Id.

165. *See, e.g.*, SQL Solutions, Inc. v. Oracle Corp., No. C-91-1079 MHP, 1991 MP, 1991 WL 626458 (N.D. Cal. Dec. 18, 1991); Wolff v. Institute of Elec. & Elecs. Eng'rs., Inc., 768 F. Supp. 66 (S.D.N.Y. 1991).

166. *Vault*, 847 F.2d at 269-70 (5th Cir. 1988). Some commentators criticize the court's reliance on the Supremacy Clause. *See, e.g.*, Gary H. Moore & J. David Hadden, *On-Line Software Distribution: New Life for 'Shrink-wrap' Licenses?*, 13 No. 4 COMPUTER LAW. 1, 5 (1996) (citing Sears, Roebuck v. Stiffel Co., 376 U.S. 225, 229 (1964)) (the authors assert that the problem with the *Vault v. Quaid* decision is that it is "not based on § 301 of the Copyright Act" and instead "relied on the 1964 Supreme Court case *Sears, Roebuck v. Stiffel Co.* for the proposition that 'when state law touches upon the area of [patent or copyright statutes], it is 'familiar doctrine' that the federal policy 'may not be set at naught, or its benefits denied' by the state law." *Id.* The authors further assert that "[t]he existence of § 301. . . makes suspect *Vault's* reliance on patent law preemption cases, since the Patent Act, unlike the Copyright Act, contains no such explicit pre-emption [sic] provision." *Id.*).

167. *See supra* note 82 and accompanying text. Constitutional preemption generally permits three types of preemption: (1) the state statute directly conflicts with a federal statute; (2) Congress may "preempt the field;" and (3) the state law stands "as an obstacle" to Congressional objectives and purposes. *See* Lemley, *supra* note 130, at 1269-70.

168. One commentator agrees that the "federal right to copy unprotected subject matter can be contractually waived by the purchaser." D.C. Toedt, *Shrinkwrap License Enforceability Issues*, 453 PLI/PAT. 613, 627 (1996). However, the state law supporting such licenses, or contracts, has the

> effect of permitting the vendor to impose restrictions on the purchaser's 'federal right

applied a constitutional analysis declaring that federal copyright law preempted the state statute and avoided basing its preemption decision on § 301 of the Copyright Act.[169] It is unclear as to the court's reason for applying Supremacy Clause preemption before § 301 copyright preemption. Perhaps the court found the breadth of § 301 copyright preemption unclear, and therefore relied on Supremacy Clause preemption.[170] Or, in the alternative, the court believed that the state cause of action passed the copyright preemption test and subsequently applied Supremacy Clause preemption.[171] In either case, the courts have consistently failed to effectively construe and rely upon § 301 copyright preemption when analyzing the enforceability of shrinkwrap licenses.

The "equivalent right" element of Copyright preemption is the crucial element in determining whether copyright law preempts elements in a contract or license.[172] Software programs and databases are subject matter covered by the Copyright Act; hence, it is this second element of copyright preemption which consumes most of the analysis when determining copyright preemption. Nimmer characterizes the equivalent right element as consisting of an "extra element" test.[173] The extra element test was applied in *Harper & Row*.[174] The court required "a qualitative difference between the asserted right and the exclusive right under the [Copyright] Act" and concluded that the state law cause of action

---

> to copy' unprotected subject matter without clear pre-purchase post-purchase notice of those restrictions. As a result, the state law is preempted by the Supremacy Clause as 'stand[ing] as an obstacle to the accomplishment and execution of the full purposes and objectives of Congress.' . . . by interfering with the federal scheme permitting the public to copy unprotectable subject matter.

Id. at 628.

169. Vault Corp. v. Quaid Software Ltd., 847 F.2d 255, 269 (5th Cir. 1988). However, the court concluded that Vault's shrinkwrap license, which prohibited the decompilation of the software, directly conflicted with § 117 of the Copyright Act and therefore preempted the statute as it "clearly 'touche[d] upon an area' of federal copyright law." *Id.* at 270.

170. Kemp, *supra* note 50, at 119. "However explicitly Congress provided for preemption in section 301, the breadth of its application is in no way clear." *Id.*

171. *Id.* at 96-97. "[A] state cause of action that passes the section 301 preemption doctrine may still be subject to preemption if it conflicts with the federal purpose, under the Supremacy Clause of the United States Constitution." *Id.*

172. Note that "equivalent right" captures both a broad or narrow interpretation of the right involved. Nimmer asserts that irrespective of the created rights restriction on copyrights, "'[t]he preemption of rights under State law is complete . . . even though the scope of exclusive rights given the work under [Copyright Act] is narrower than the scope of common law rights in the work might have been.'" NIMMER, § 1.01[B][1], *supra* note 85, at 1-11 (citing H.R. Rep. No. 101-514, 101st Cong. 2d Sess. 21 (1990)). Therefore, in *Vault*, for example, where the software developer sought to restrict rights to its software package, such a restriction constitutes an "equivalent" right under the Copyright Act and should be preempted.

173. *See supra* note 85 and accompanying text.

174. Harper & Row, Publishers, Inc. v. Nation Enters., 723 F.2d 195 (2d Cir. 1983), *rev'd on other grounds*, 471 U.S. 539 (1985).

did not "establish qualitatively different conduct on the part of the infringing party, nor a fundamental nonequivalence between the state and federal rights implicated."[175]

This extra element test applies to breach of contract claims such as those asserted by vendors of shrinkwrap licenses. If the user breaches the terms of the license, a breach of contract claim is filed and a court must establish whether such a claim is "qualitatively" different from a claim in copyright infringement.[176] For example, in *National Car Rental System, Inc. v. Computer Associates International, Inc.*, a software developer's licensing agreement restricted the use of its computer programs to the processing of the purchaser's data.[177] The purchaser subcontracted its data processing department to a computer consulting company. All parties agreed that the consulting company could use the licensed software to operate the purchaser's computer systems.[178] The software developer asserted the purchaser made an unauthorized copy of the programs by allowing the computer consulting company to process third-party data.[179] The Eighth Circuit held that the breach of contract claim was not preempted because "the alleged contractual restriction on [the purchaser's] use of the licensed programs constitutes an extra element in addition to the copyright rights making this cause of action qualitatively different from an action for copyright."[180] Thus, *National Car* "stands for the proposition that a breach of contract action is qualitatively different from a claim on copyright infringement when the right or obligation being enforced would not exist *but for the parties' agreement*."[181]

Unlike the licensing agreement in *National Car*, however, shrinkwrap licenses are not negotiated contracts. Under a shrinkwrap license, the parties do

---

175. *Id.* at 201.

176. *See supra* note 86 and accompanying text. Courts have wrestled with this application in breach of contract claims. *See* Wolff v. Inst. of Elec. & Elecs. Eng'rs, Inc., 768 F. Supp. 66 (S.D.N.Y. 1991) (holding that where the act which caused the breach of contract is the same act which would trigger a claim in copyright infringement, the breach of contract claim is preempted). *But see* National Car Rental Sys., Inc. v. Computer Assoc. Int'l, Inc., 991 F.2d 426 (8th Cir. 1993) (reversing the district courts' finding of preemption).

177. *National Car*, 991 F.2d at 427-28.

178. *Id.* at 428.

179. *Id.*

180. *Id.* at 431. The court further concluded that "the contractual restriction on use of the programs constitutes an additional element making this cause of action not equivalent to a copyright action." *Id.* at 432.

181. Maureen A. O'Rourke, *Drawing the Boundary Between Copyright and Contract: Copyright Preemption of Software License Terms*, 45 DUKE L.J. 479, 521 (1995) (emphasis added). Professor O'Rourke suggests that the bargaining which occurs in negotiated contracts provides the extra element required to satisfy the "equivalent" rights test of § 301. *Id.* at 522. Yet, as far as standard form contracts are concerned, as the U.C.C. Draft indicates, the mutual assent through conduct may also constitute the extra element necessary to avoid preemption. *Id.* at 530-31; *see also* U.C.C. Draft, *supra* note 141, at § 2B-308.

not confer and come to an agreement on the terms of the contract limiting usage rights. Thus, this lack of negotiation, or bargaining, does not satisfy the extra element test as applied in *National Car*. Therefore, the extra element of "party agreement," established in *National Car*, forecloses similar application to shrinkwrap licenses.[182]

Instead, Judge Easterbrook considered and relied upon market conditions to determine shrinkwrap license enforceability under the rubric of a § 301 preemption analysis.[183] ProCD asserted that contracts of "private ordering" were "essential to the efficient functioning of markets."[184] Easterbrook focused on the value of ProCD's database to the marketplace and described the theory and justification of price discrimination between commercial and non-commercial products.[185] Yet, unlike *Vault v. Quaid*, which omitted a section 301 "equivalent rights" analysis, Judge Easterbrook applied the equivalent rights test in his preemption analysis. The court in *ProCD* held that rights created by contract, presumably even if not negotiated, are not "equivalent to any of the exclusive rights within the general scope of copyright."[186]

> Copyright law forbids duplication, public performance, and so on, unless the person wishing to copy or perform the work gets permission; silence means a ban on copying. A copyright is a right against the world. Contracts, by contrast, generally affect only their parties; strangers may do as they please, so contracts do not create "exclusive rights." Someone who found a copy of SelectPhone™ on the street would not be affected by the shrinkwrap license—though the federal copyright laws of their own force would limit the finder's ability to copy or transmit the application program.[187]

The Seventh Circuit did not apply the extra element test, but instead compared copyright preemption to another federal preemption clause involving airline carriers.[188] The court borrowed the reasoning in *American Airlines, Inc. v. Wolens* which concluded that "[t]erms and conditions offered by contract

---

182. O'Rourke, *supra* note 179, at 522-23. Yet, when such limiting provisions are found in non-negotiated, standard form licenses, such as shrinkwrap licenses, "the argument for nonpreemption of breach of contract actions is more tenuous." However, Professor O'Rourke concedes that in either case, the court should "make a deeper inquiry into market conditions before making the preemption decision." *Id.*

183. ProCD, Inc. v. Zeidenberg, 86 F.3d 1447, 1449-50 (7th Cir. 1996).

184. *Id.* at 1455.

185. *Id.* at 1449-50. Judge Easterbrook likened computer purchases to insurance purchases, airline ticket purchases and concert ticket purchases characterizing them as "[t]ransaction[s] . . . [where the] money precedes the communication of detailed terms." *Id.* at 1451.

186. *Id.* at 1455.

187. *Id.* at 1454.

188. *Id.* (quoting 49 U.S.C. App. §1305(a)(1) (1988)) ("[This] federal statute preempts any state 'law, rule, regulation, standard, or other provision . . . relating to rates, routes, or services of any air carrier.").

reflect private ordering, essential to the efficient functioning of markets."[189] Under this reasoning, a shrinkwrap license may reflect "private ordering" and should prevail over national directives in order to support efficient marketing systems. However, the court refined its ruling, maintaining that merely because a claim might fall under the auspices of a "contract," it did not automatically avoid § 301 preemption analysis.[190] Some contracts could interfere with the national objectives of copyright law,[191] but in *ProCD*, the shrinkwrap license did not interfere with copyright objectives.[192] "[W]hether a particular license is generous or restrictive, a simple two-party contract is not 'equivalent to any of the exclusive rights within the general scope of copyright' and therefore may be enforced."[193]

ProCD marks a new era of shrinkwrap license enforcement. Certainly, this decision follows the course which the U.C.C. Drafting Committee has charted.[194] The court's rational, however, tends to interfere with copyright policy. When a shrinkwrap license grants or restricts rights which affect the public domain, copyright law should prevail.[195] A court should assess the effect shrinkwrap license enforcement has upon the balance between the author's right and the public's right to access information.[196] As such, the U.C.C. Drafting Committee should also recognize this balance and invalidate a shrinkwrap license if it purports to afford or restrict rights "equivalent" to the Copyright Act.[197]

The shrinkwrap license serves as a crutch to database protection. The database developer seeks to protect against a commercial invasion of the work

---

189. *Id.* at 1455 (quoting American Airlines, Inc. v. Wolens, 115 S. Ct. 817, 824-25 (1995)).

190. *Id.* The court pointed to *Wolens* and *National Car* for support in this conclusion: "National Car Rental likewise recognizes the possibility that some applications of the law of contract could interfere with the attainment of national objectives and therefore come within the domain of § 301(a)." *Id.*

191. *Id.* The court's comparative analysis found that § 301(a) and § 1305(a)(1) provide a similar function: they "prevent[] states from substituting their own regulatory systems for those of the national government." *Id.*

192. *Id.*

193. *Id.*

194. The next version of the U.C.C. Draft is sure to cite to the *ProCD* decision in its Reporter's Notes section to support the Committee's position of the viability and enforceability of shrinkwrap licenses.

195. 17 U.S.C. § 301 (1994).

196. *See* Lemley, *supra* note 130. The author contends that the federal intellectual property law is "sufficiently comprehensive" and "that no two parties should be allowed to alter or avoid some aspect of intellectual property law where the result is to disadvantage others who are not a party to the contract." *Id.* at 1291.

197. *Id.* The author recommends a limitation for the U.C.C. Draft Mass Market License Section 2B-308: "a term does not become part of the license if the term . . . creates an obligation or imposes a limitation on the licensee that is inconsistent with federal intellectual property law, or that deprives the licensee of a right or privilege granted the licensee under federal intellectual property law." *Id.* at 1292.

which would lower development costs of the subsequent developer.[198] In essence, the database developer seeks an unfair competition law to protect the initial investment. The shrinkwrap license presumably serves as this protection. However, while it may protect the developer from commercial exploitation, it also limits endusers' rights—a non-commercial user.

What is actually sought is the protection of the first compiler's information—a "compulsory license."[199] A compulsory license could provide the "appropriate balance[] between securing the commercial value of low authorship compilations on the one hand and promoting creation of and access to a wide variety of informational works on the other."[200] Under a compulsory licensing scheme, the subsequent developer, seeking to use the initial developer's work, would request a license from the initial developer prior to using any information in the initial developer's database. The subsequent developer would pay a fee to the initial developer to use the information. Such a licensing scheme would provide sufficient protection to the database developer with significant investment in data compilation yet still afford the public access to derivative databases.[201]

Without assurance that the database is protected by copyright law or shrinkwrap licenses, the database developer might be reluctant to create comprehensive works.[202] This conflicts with the policy of motivating authors to

---

198. Hayden, *supra* note 30, at 240. "A developer's real concern is that a competitor will take the product of his efforts and use it to develop a competing product that, because of its reduced development costs, can be sold at a much lower price. Expansive protection . . . would simply prevent unjust enrichment by subsequent developers." *Id.*

199. Ginsburg, *supra* note 113, at 1917. Professor Ginsburg advocates amending copyright law to "recognize the first compiler's rights over derivative versions of informational works, but would qualify those rights by compelling information providers to license rights to produce derivative compilations." *Id.* at 1916.

200. *Id.* Professor Ginsburg provides the rationale for compulsory licensing supporting the balance between incentives and public access through an example of a database:

> Maximum incentives might be afforded to data collection were the first gatherer of the dataset to receive exclusive control over any recombinations that might be made of information contained within the compilation; nonetheless, vesting this control in a single compiler would cut off public access to new informational works that could be generated from the data, but that the compiler declines to license. It is not a sufficient answer to reply that, once the copyright expires, anyone may freely revise the data; seventy-five years is a long time to suspend the public interest in the new combinations of gathered information.

201. *Id.*

202. Robert Gorman, *Copyright Protection for the Collection and Representation of Facts*, 76 HARV. L. REV. 1569 (1962); Heald, *supra* note 99, at 152 (citing William M. Landes & Richard A. Posner, *An Economic Analysis of Copyright Law*, 18 J. LEGAL STUD. 325 (1989); Jessica Litman, *The Public Domain*, 39 EMORY L.J. 965 (1990); Ginsburg, *supra* note 113); John S. Wiley Jr., *Copyright at the School of Patent*, 58 U. CHI. L. REV. 119 (1991).

create works for the public good.[203] Thus, where does the database developer turn to get the assurance needed to motivate her to create databases the public seeks and which ultimately "promote the Progress of Science and useful Arts"?[204]

B. In Search of a Sui Generis Right for Database Protection

1. An International Solution to the Database Protection Dilemma.— Database protection through shrinkwrap licenses leaves the developer exposed to potential copyright preemption.[205] Although the Seventh Circuit enforced a shrinkwrap license, if the other courts followed suit, the United States would be among the few countries in the Western world enforcing such licenses.[206] Other countries either directly refuse to enforce shrinkwrap licenses or place restrictions on the effect of such licenses.[207] However, restrictions lend themselves to uncertain database protection as well—the precise dilemma we face today. Hence, the database developer maintains limited legal protection. To improve database protection, some advocate the creation of a *sui generis* right.[208]

To encourage investment in database development, the European Community (EC) recently adopted the EC Database Directive creating a *sui generis* right protecting a database from "unauthorized extraction or re-utilization" of database contents.[209] This directive creates a protective right outside the realm of member

---

203. Heald, *supra* note 99, at 152-55.

204. U.S. CONST. art. I, § 8, cl. 8.

205. *But c.f.* ProCD, Inc. v. Zeidenberg, 86 F.3d 1447 (7th Cir. 1996) (court enforces shrinkwrap license and holds copyright preemption does not prevail).

206. *See* Lemley, *supra* note 130, at 1253 n.54. The author asserts that only five countries enforce shrinkwrap licenses: Austria, Dominican Republic, Hong Kong, Korea, and Malaysia. *Id.*

207. *Id.* at 1253 n.53. Some of these foreign countries will enforce shrinkwrap licenses if certain conditions are met. *Id.*

208. For articles supporting the creation of *sui generis* right for database protection, see generally, Gross, note 52 (author argues that there is a need for sui generis legislation for software programs); Hayden, *supra* note 30 (argues current copyright law is ineffective for database protection and asserts that a *sui generis* right for such protection is necessary); Howell, *supra* note 68 (author supports development of supplemental legislation to protect databases against piracy; supports specifically the passage of the legislation which criminalizes piracy: the proposed Criminal Copyright Improvement Act of 1995); Teddy C. Kim, Note, *Taming the Electronic Frontier: Software Copyright Protection in the Wake of* United States v. Lamacchia, 80 MINN. L. REV. 1255 (1996) (arguing that criminal sanctions for piracy is "unworkable and undesirable" and that Congress should revise the current Copyright Act in order to protect digital technologies). *But see* Ginsburg, *supra* note 113 (Professor Ginsburg advocates the amendment of the Copyright Act in support of compulsory licensing for derivative works); Litman, *supra* note 30 (author is leery of advancing the need for a *sui generis* right under the Commerce Clause for protection of facts as it would potentially limit public access to this information).

209. Paul Durdik, *Ancient Debate, New Technology: The European Community Moves to Protect Computer Databases*, 12 B.U. INT'L L.J. 153, 165 (1994) [hereinafter EC Database

countries' copyright law, that prevents "unfair use of database contents where significant cost and efforts were expended to create it."[210] Under *Feist*, such cost and labor-based protection does not exist for a U.S. compilation. Perhaps in recognition of the limitations *Feist* imposes on U.S. technology companies, the EC Database Directive does not reciprocate protection unless the contracting country has similar protection for its national databases.[211] This leaves United States technology companies vulnerable in the European Community.

In September 1995, the European Community submitted a discussion paper to the Committee of Experts of the World Intellectual Property Organization ("WIPO") on the *sui generis* right for database protection. Six months later, the European Community submitted their proposal for "international harmonization of the *sui generis* protection of databases" to WIPO.[212] The WIPO Database Treaty's preamble asserts that the contracting parties

> [r]ecognizing that databases are a vital element in the development of a global information infrastructure and an essential tool for promoting economic, cultural and technological advancement, [r]ecognizing that the making of databases requires the *investment of considerable human, technical and financial resources* but that such databases can be copied or accessed at a fraction of the cost needed to design them independently, [and] [d]esiring to establish a new form of protection for databases by granting rights adequate to enable the makers of databases to recover the investment they have made in their databases and by providing international protection in a manner as effective and uniform as possible. . . .[213]

The very preamble of the Database Treaty rejects *Feist* by resurrecting the "sweat of the brow" theory as does the EC Database Directive. Article 1 of the Database Treaty further asserts that

> [the] [c]ontracting [p]arties shall protect any database that represents a substantial investment in the collection, assembly, verification, organization or presentation of the contents of the database. . . . The protection granted under this Treaty shall be provided irrespective of any

---

Directive]. The author points to the compulsory licensing provision which the EC Database Directive includes to avoid the monopolization of information compiled in these databases. *Id.* at 166.

210. Barry D. Weiss, *Barbed Wires and Branding in Cyberspace: The Future of Copyright Protection*, 450 PLI/PAT. 397, 404 (1996).

211. *Id.*

212. *Basic Proposal for the Substantive Provisions of the Treaty on Intellectual Property in Respect of Databases to be Considered by the Diplomatic Conference*, World Intellectual Property Organization (WIPO), CRNR/DC/6, 6 (Aug. 30, 1996). <http://www.loc.gov./copyright/wipo6.html> [hereinafter WIPO Draft Database Treaty]. The European Community has been instrumental in the draft of the Database Treaty. *Id.* at 5.

213. *Id.* at Preamble (emphasis added).

> protection provided for a database or its contents by copyright or by other rights granted by Contracting Parties in their national legislation.[214]

An additional provision of the Database Treaty allows "[t]he maker of a database . . . [to] have the right to authorize or prohibit the extraction or utilization of its contents."[215] Thus, the very extraction of data from a database, which took considerable human and technical resources to develop, would violate this treaty. Yet, even WIPO recognizes the treaty's potential conflict with copyright laws.

> Article 12 [of the Database Treaty] deals with the relationship between the protection accorded under the proposed Treaty and existing or future rights and obligations. The protection granted under the proposed Treaty shall leave intact and shall in no way affect any "conventional" rights in the database or its contents.[216]

The European Community has led the way in database protection reform. Now, the United States and other countries must decide if and how they intend to operate in the global technology industry. U.S. Internet and online service providers oppose the support of the WIPO Database Treaty because database protection issues have not been fully discussed domestically.[217] Others contend that it is premature to push an untested area of law into the international arena without having experienced the ramifications of database protection domestically.[218] WIPO agreed. In December 1996, WIPO held a Diplomatic

---

214. *Id.* art. I.

215. *Id.* art. III. The term "extraction" is defined as "the permanent or temporary transfer of all or a substantial part of the contents of a database to another medium by any means or in any form." *Id.* art. II.

216. *Id.* art. XII (notes to Article 12). Article 12 reads:

> The protection accorded under this Treaty shall be without prejudice to any other rights in, or obligations with respect to, a database or its contents, including laws in respect of copyright, rights related to copyright, patent, trademark, design rights, antitrust or competition, trade secrets, data protection and privacy, access to public documents and the law of contract.

Id.

217. *Copyrights: Clinton WIPO Treaty Proposals Meet with Significant Opposition*, BNA PAT., TRADEMARK & COPYRIGHT L. DAILY, Sept. 23, 1996, at D2. Senator Orrin G. Hatch: "Congress will not wish to be in the position of having its hands tied by international developments on the basis of proposed legislation that has stalled precisely because it contains so many unresolved issues." *Id.* Professor Pamela Samuelson, professor of law and information management at the University of California at Berkeley, contends that the "administration's international proposals are aimed at partially preempting domestic discussion on contentious copyright issues." *Id.*

218. *Id.* However, others have noted the EC Database Directive requires reciprocity "which means that any producer of databases not located in a European Union country would not be given protection unless the producer's home nation had comparable legislation in place." *Id.* The EC Database Directive generally provides for Member States to enact such legislation by 1998. EC

Conference in Geneva to discuss the WIPO Database Treaty proposal.[219] The Conference, proved that many countries "weren't ready to address" extending protection to the data in a database.[220] The Conference concluded that further study of the proposed treaty was needed.[221] Hence, even the international community has come full circle demonstrating their uncertainty as to the best approach for electronic database protection.

2. A U.S. Proposal for Database Protection: The Intellectual Property Antipiracy Act.—Amending the Copyright Act for increased database protection may be prohibitive. *Feist* may have tied Congressional hands when it held that originality was not only a statutory requirement for determining copyrightability, but a Constitutional one as well.[222] Although the Copyright Clause charges Congress with "increas[ing] public welfare by providing incentives for creation," the decision in *Feist* effectively "forbids [the] protection of labor, research, or industrious collection" thereby dampening the economic incentive to develop commercially viable databases.[223]

One constitutional option remains. Congress can create *sui generis* database protection under the Commerce Clause.[224] Since the Court prohibited protection of low authorship works under the Copyright Clause, evading *Feist* by resurrecting the sweat of the brow theory under the Commerce Clause might not gather much support.[225] The Copyright Clause mandates the protection of

---

Database Directive, *supra* note 207, at art. 32, 33. However, Member States can apply for three to nine year delays in enacting such legislation. *Id.* This demonstrates some reluctance on the part of the very countries that proposed this international treaty evidencing uncertainty as to the best method of database reform.

219. *Copyright Accords Guard Against Software Piracy*, WALL ST. J., Dec. 23, 1996, at B6.

220. *Id.*

221. *Lehman Calls WIPO Conference a Success; Attention now Turns to Treaty Ratification*, BNA PATENT, TRADEMARK & COPYRIGHT LAW DAILY, Jan. 22, 1997, at D2. In asserting the need to further study database protection, Lehman noted that "it took six years to study the copyright treaties." *Id.*

222. Feist Publications, Inc. v. Rural Tele. Serv. Co., Inc., 499 U.S. 340, 351 (1991). Professor Heald suggests that "the familiar white pages are unprotected and Congress apparently can do nothing to render them protectable short of initiating a constitutional amendment." Heald, *supra* note 99, at 144.

223. *Id.* at 158. Professor Heald further contends that the Supreme Court in *Feist* clearly "locks Congress into . . . [providing] virtually no protection . . . [for] labor and research because the cost of such protection is simply too high." *Id.* at 159.

224. *Id.* at 168. Should Congress respond to *Feist* by protecting works through a "sweat of the brow" theory, Congress would have to "enact a broader unfair competition statute prohibiting the copying of any 'industrious collection.'" *Id.*

225. *Id.* at 175. Professor Heald also points out that enacting such legislation under the Commerce Clause could "pose[] serious difficulties because the Court made it clear in *Feist* that such protection would directly conflict with the purpose of the [Copyright] Clause. According to the Court, the clause not only tells us what may be protected, but what must remain in the public domain." *Id.* at 172.

original works of authorship; circumventing this originality requirement by enacting legislation under the Commerce Clause may be unconstitutional.[226]

Nonetheless, in May 1996, Representative Moorhead introduced the Database Investment and Intellectual Property Antipiracy Act of 1996 which proposed an amendment to Title 15 of the United States Code.[227] The Database Investment Act prohibited the "extract[ion], use or reuse of all or a substantial part . . . of the contents of a database . . . that conflicts with the database owner's normal exploitation of the database or adversely affects the actual or potential market for the database. . . ."[228] The Act further proscribed the use of the extracted data in a market which "directly or indirectly" competed in the market from which the data was extracted.[229] This act also provided an abridged version of fair use: "a lawful user of a database made available to the public or placed in commercial use is not prohibited from extracting, using or reusing insubstantial parts of its contents, qualitatively or quantitatively, for any purposes whatsoever."[230] A limiting aspect of this section concerns the lack of fair use rights to libraries or the educational community.[231] The boldest provision, however, appears in section three of the Database Investment Act.

> A database is subject to the Act if it is the result of a qualitatively or quantitatively *substantial investment of human, technical, financial or other resources in the collection, assembly*, verification, organization or presentation of the database contents, and (i) the database is used or reused in commerce; or (ii) the database owner intends to use or reuse the database in commerce.[232]

This adaptation advocated the protection of databases as a result of "industrious collection" in direct contravention to the *Feist* decision.[233]

---

226. Professor Heald analogizes such action under the Bankruptcy Clause. In 1980, Congress attempted to protect "employees of the bankrupt . . . Rock Island estate, and not [extend such protection] to other railroad bankruptcies" which violated the "uniform" requirement of the Bankruptcy Clause. A unanimous Court responded that Congress could not enact such legislation under the Commerce Clause to bypass the Bankruptcy Clause's requirement of uniformity. *Id.*

227. H.R. 3531, 104th Cong., 2d Sess. (1996) [hereinafter Database Investment Act]. This bill is also available on the Internet at gopher://wgate.house.gov. The official title of this bill reads: A bill to amend title 15, United States Code, to promote investment and prevent intellectual property piracy with respect to databases.

228. Database Investment Act, *supra* note 225, at § 4(a)(1).

229. *Id.* § 4(b)(1).

230. *Id.* § 5(a).

231. See 17 U.S.C. §§ 107, 108 (1994) for fair use rights extended to libraries and the educational community. *See supra* notes 75 and accompanying text.

232. Database Investment Act, *supra* note 225, at § 3(a) (emphasis added).

233. However, in a section entitled "Relationship to Other Laws," this act expressly acknowledged *Feist*, copyright law, and shrinkwrap licenses. *Id.* § 9.

> The remedies against violations hereunder shall be without prejudice to any remedies under any copyright that may subsist in the database, any contents of the database, or

The bill could have threatened the delicate balance of rights in the author's work with its availability to the public. The incentive to create would prevail as the database developer would then have assurances that the information in the database would not be misappropriated in a commercial fashion to compete with their database creation.[234] However, the same Act failed to adequately protect public access to the information in the developer's database—how much information could a student copy for use in a research project?[235]

The Database Investment Act died on the house floor at the end of the 104th session.[236] The reluctance to extend protection to facts, as in Geneva, prevails in Congress. Such a monopolistic approach to data is undesirable. Even if Congress drafts a bill which does not threaten the balance between the rights of the author and the availability to the public, a Constitutional question remains: may Congress constitutionally overrule the *Feist* decision by enacting a statute under the Commerce Clause? The answer remains tenuous at best.

V. AN ANALYSIS OF ALTERNATIVE APPROACHES TO DATABASE PROTECTION

A. ALTERNATIVE I: The Shrinkwrap License

Shrinkwrap licenses perpetuate a false sense of reliance in database protection. Shrinkwrap licenses furnish database developers with a security blanket because they project an *illusion* of protection. Conversely, *ProCD* and the U.C.C. Draft, which uphold shrinkwrap license enforceability, provide additional assurances to the developer. Why then this false sense of security? Copyright preemption. Section 301 of the Copyright Act explicitly preempts attempts to create "equivalent" rights to copyright law.[237]

The U.C.C. Draft recognized that shrinkwrap licenses tend to expand or restrict rights that conflict with copyright law. Consequently, the drafters are currently debating a proposed amendment to U.C.C. 2B-308 that prevents a term inconsistent with copyright law from becoming part of the contract.[238]

> the selection, coordination or arrangement of such contents. . . . Nothing in this Act shall restrict the rights of parties freely to enter into licenses or any other contracts with respect to databases or their contents.

Id. § 9(a) & (b).

234. *Id.* § 4.

235. *Id.* § 5. The Database Investment Act, for example, failed to provide fair use provisions for educational purposes. *Id.* Further, the Act acknowledged the "rights of parties freely to enter into licenses or any other contracts with respect to databases or their contents." *Id.* § 9(b). This is the type of legislation Professor Litman feared: protection of database rights at the expense of the public domain. Litman, *supra* note 200.

236. Lisa H. Greene & Steven J. Rizzi, *Database Protection Developments: Proposals Stall in the United States and WIPO*, 9 No. 1 J. PROPRIETARY RTS. 2, 5 (1997).

237. 17 U.S.C. § 301 (1994).

238. *Copyright*, *supra* note 149, at D3. *See also* Lemley, *supra* note 130, at 1292 (proposing a similar amendment under an older draft of the mass market license section: "a term does not

Yet, even this provision does not remedy the fundamental issue of copyright preemption of shrinkwrap licenses. Congress explicitly included a preemption section in the Copyright Act instead of relying solely upon the Supremacy Clause. This further demonstrates Congress' intent to protect copyright laws without succumbing to agreements which arbitrarily expanded an owner's rights in a work. Therefore, when properly analyzed and applied, copyright preemption weakens the shrinkwrap license approach as a viable solution to database protection.

B. ALTERNATIVE II: A Sui Generis Right for Database Protection

A more viable alternative might include *sui generis* database protection under the Commerce Clause. This option would preserve copyright policies—creating incentives for database development while meeting public demands for access to large repositories of information. Moreover, this option supports the assertion that if the work would otherwise go uncreated, there is a need to protect it.[239] Additionally, given the EC Database Directive, such works may go unprotected if reciprocal protection is not provided nationally.[240]

The issue, however, is whether such legislation can survive judicial scrutiny. The Commerce Clause grants to Congress the power "[t]o regulate Commerce with foreign Nations, and among the several States. . . ."[241] In response to international activities, Congress may assert its mandated right to create legislation to "regulate Commerce with foreign Nations."[242] The EC Directive calls for reciprocal legislation protecting database investments.[243] The EC, as well as database developers, seek legislation which would prohibit substantial extraction or re-use of the database. This would prevent a subsequent developer from misappropriating the work and marketing it in a similar market thereby adversely affecting the original developer's market potential. Without such protection, EC technology companies will be less likely to participate in the U.S. market while U.S. technology companies lose their viability in the global economy. This lack of availability and access to global electronic databases also disturbs the founders desire to provide the public with creative works to stimulate and foster progress. Hence, it is likely that *sui generis* database protection, responding to international commerce concerns, could survive judicial scrutiny.

This option provides two avenues for developing such protection: copyright laws or the protectionary provisions under *sui generis* legislation. Providing two avenues of protection to the technology industry is not a new development. For

---

become part of the license if the term: . . . creates an obligation or imposes a limitation on the licensee that is inconsistent with federal intellectual property law, or that deprives the licensee of a right or privilege granted the licensee under federal intellectual property law.").

239. Heald, *supra* note 99, at 152-55.

240. *See supra* note 209 and accompanying text.

241. U.S. CONST. art. I, § 8, cl. 3 (the Commerce Clause).

242. *Id.*

243. *See supra* note 209 and accompanying text.

example, software vendors may opt for protection pursuant to the Copyright Act, patent laws[244] or trade secret laws.[245] Although providing protection across two areas of federal law might appear unappealing, such direction has been taken in the past and may offer a viable alternative for the future.

Congress should consider including compulsory licensing as part of this new form of protection.[246] The database developer seeks the protection of his exhaustive investment. A compulsory licensing provision could provide for protection of the first compilation of facts, but require subsequent developers to procure a license from the initial developer for commercial derivative databases.[247] This option permits facts to remain in the public domain while promoting the development of future works.

Additionally, the *sui generis* right should elaborate on fair use rights to the work. Such a provision should extend fair use to the education and research communities. Further, a fair use provision could provide limitations on the abuse of database extraction. Such a provision should allow a lawful user of a commercial database to extract or reuse insubstantial amounts of data from the database. If the extracted data serves educational needs (e.g., library use, classroom copies or research use), and is not used to compete commercially with the database, such use should be deemed inconsequential. A fair use provision is best applicable "when the objectives of the . . . system would be frustrated rather than furthered by a finding of [misappropriation]."[248]

Given that the recent Database Investment Act died on the House floor, considerable study and discussion is needed to evaluate this sensitive area of law. Congress may again need to turn to a committee to assess the needs of both the technology industry and the general public to establish proper reform. In 1974, Congress established CONTU to address copyright issues with respect to technology which resulted in effective recommendations.[249] Perhaps it is time for Congress to institute "CONTU II" to address similar protectionary needs in the database industry. Such a commission should assess the viability of developing a new *sui generis* right, with elements of fair use as well as preventive measures for commercial misappropriation. This would protect developers' rights in the works, the public domain, and the common enduser/consumer of electronic goods. This alternative could provide effective protection domestically and

---

244. *See generally* Wayne M. Kennard, *Obtaining and Litigating Software Patents and Protecting Software on the Internet*, 444 PLI/PAT. 275 (1996).

245. *See generally* Victoria A. Cundiff, *Protecting Computer Software as a Trade Secret*, 444 PLI/PAT. 7 (1996).

246. *See* Ginsburg, *supra* note 113, at 1925. A similar approach is advocated through amendment of the Copyright Act to allow for compulsory licensure of low authorship works. *Id.* at 1927.

247. *Id.* at 1924-27. "The effect of compulsory license is to grant open access to the covered material, subject to an obligation to pay the owner for the use. Compulsory licensing substitutes compensation for control over the copyrighted work." *Id.* at 1925.

248. Denicola, *supra* note 115, at 524.

249. *See generally* CONTU Report, *supra* note 58.

internationally without offending the fundamental principles of the Copyright Clause. The creation of a *sui generis* right offers the best viable alternative to database protection.

CONCLUSION

Database protection under copyright law has deteriorated. Today, users seek information that is presented in a more efficient manner; this has fueled the demand for access to databases. Copyright law has had difficulty keeping pace. Consequently, shrinkwrap licenses have increased as the prevalent method for database protection. Yet, given the copyright preemption provision, such methods are ineffective.

This Note supports the creation of a *sui generis* right for database protection. While such an approach develops and endures refinement and debate, database developers, in the interim, will have to rely upon current copyright laws. In so doing, developers could advocate fair use of their database to protect against commercial inequities or misappropriation.

In repairing database protection, careful consideration should be given to the underlying policies of copyright law that stem from the framers of the Constitution.[250] Congress has a constitutional mandate to provide incentives to authors to create works for the benefit of the public.[251] Moreover, reform must effectively balance the protection of copyrighted works with the public interest. This necessarily includes an assessment of this balance in a global economy. The European Community has recognized the need and has aggressively pursued a solution. The Executive Branch also recognizes the need to ensure better copyright protection of online materials: "[t]he [Intellectual Property and the National Information Infrastructure] report suggests that the Internet will not flourish if significant protection against theft and copyright abuse is not offered."[252] Similarly, the database industry will not flourish if proper protection is denied. To encourage database development, an aggressive study and resulting legislation is needed in order to continue to further the "Progress of Science and useful Arts."[253]

---

250. The framers expressly granted to Congress the power "[t]o promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries." U.S. CONST. art. I, § 8, cl. 8.

251. Denicola, *supra* note 115, at 519. "[T]his mandate embraces a tension between economic incentives to produce, which are provided through the recognition of exclusive rights, and unfettered access to creative works in order to maximize dissemination and to enable past inventiveness to serve as the foundation for future contributions." *Id.*

252. *ProCD, Inc. v. Zeidenberg*, 908 F. Supp. 640, 647 (W.D. Wisc. 1996) (citing Guy Alvarez, *New Legal Issues on the Net*, AM. LAW. 28, 29 (Dec. Supp. 1995)).

253. U.S. CONST. art. I, § 8, cl. 8.

APPENDIX: U.C.C. Draft § 2B-308. Mass Market Licenses.

(a) Except as otherwise provided in section and Section 2B-309, a party adopts the terms of a mass-market license if the party agrees or manifests assent to the mass market license before or in connection with the initial use of or access to the information.

(b) Terms adopted under subsection (a) include all of the terms of the license without regard to the knowledge or understanding of individual terms by the party assenting to the form. However, except as otherwise provided in this section, a term does not become part of the contract if the term creates an obligation or imposes a limitation which:

(1) the party proposing the form should know would cause an ordinary and reasonable person acquiring this type of information and receiving the form to refuse the license if that party knew that the license contained the particular term; or

(2) conflicts with the previously negotiated terms of agreement.

(c) A term excluded under subsection (b) is part of the contract if the party that did not prepare the form manifests assent to the term or if, under the circumstances, the limitation or obligation in the term was clearly disclosed to the party before it agreed or manifested assent to the mass market license.

(d) A term of a mass market license which is unenforceable for failure to satisfy a *175 requirement of another provision of this article, such as a provision that expressly requires use of conspicuous language or manifested assent to the term, is not part of the contract.

(e) A mass-market license must be interpreted whenever reasonable as treating in a similar fashion all parties situated similarly without regard to their knowledge or understanding of the terms of the record.

(f) A term that states a limitation that would be placed on the party by copyright or patent law in the absence of the term does not within subsection (b)(1).

SEMINOLE TRIBE AND SUPERFUND: A FEDERALISM GAMBLE

BARRY L. LOFTUS*

The King can do no wrong[1]
Nemo est supra leges[2]

INTRODUCTION

A casualty of the Supreme Court's recurrent battle over federalism—the constitutional relationship between the federal and state governments[3]—is the balance between state sovereignty and state accountability. Specifically, private citizens no longer may pursue many federal rights of action against unconsenting states in federal courts. As a stark example of this power shift, private parties seeking just compensation from states have lost significant rights under a cornerstone of the nation's environmental regulatory enforcement scheme, the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA or Superfund).[4]

In *Seminole Tribe of Florida v. Florida*,[5] the third of four major federalism case for the Rehnquist Court in which states' rights were enhanced at the expense of Congress,[6] the Court held that Congress does not have authority under the

---

* J.D. Candidate, 1998, Indiana University School of Law—Indianapolis; B.A., 1985, Marquette University. The author thanks his wife, Kimberly G. Loftus, for her suggestions and support.

1. Ancient maxim. *See* Louis L. Jaffe, *Suits Against Governments and Officers: Sovereign Immunity*, 77 HARV. L. REV. 1, 3-5 (1963) (The doctrine of sovereign immunity first emerged in the reign of Edward I (1272-1307) based on the maxim "The King can do no wrong." Jaffe suggests that the statement really means the King is not allowed to do wrong. The British monarchs, desiring to extend justice to subjects, got around this problem by making the King's officers amenable to suit.)

2. "No one is above the law." BLACK'S LAW DICTIONARY 1038 (6th ed. 1990).

3. *Id.* at 612. *See* Erwin Chemerinsky, *Restricting Federal Court Jurisdiction*, TRIAL, July 1996, at 18 (stating that "[t]he most important changes in constitutional law in the 1990s have been in the area of federalism.").

4. 42 U.S.C. §§ 9601-9675 (1994).

5. 116 S. Ct. 1114 (1996).

6. The other three cases are *New York v. United States*, 505 U.S. 144, 162 (1992) (Congress may, under its power to regulate private activity under the Commerce Clause, offer states the choice of regulating the activity according to federal standards or of having state law pre-empted by federal regulation.); *United States v. Lopez*, 514 U.S. 549, 561 (1995) (provision of Gun-Free School Zones Act, 18 U.S.C. § 922(q) (1994), which made it illegal to possess firearms in a school zone, was an unconstitutional expression of Congress' power under the Commerce Clause because the regulated activity did not substantially affect interstate commerce); *City of Boerne v. Flores*, 117 S. Ct. 2157, 2172 (1997) (Religious Freedom Restoration Act, 42 U.S.C. §§ 2000bb-1 to -4 (1994) exceeds Congress' power under the Enforcement Clause of the Fourteenth Amendment, contradicting principles necessary to maintain federal-state balance).

Commerce Clause[7] to grant private parties federal causes of action against unconsenting states.[8] While the specific statute invalidated in the case concerned Indian gaming regulations, the Court specifically overruled one of its CERCLA cases that recognized private contribution and cost-recovery actions against states.[9]

The decision thwarts the twin policy objectives of CERCLA of making polluters pay for their damage and encouraging efficient, voluntary cleanups.[10] States have broad authority under CERCLA to assist in cleanups, but states are also liable for damage caused as a result of their ownership and operation of sites. CERCLA litigation often features scores of liable parties, including state governments, battling each other to recover cleanup costs. The *Seminole Tribe* ruling allows states to use the Eleventh Amendment[11] of the Constitution to bar private suits in federal courts arising from laws rooted in the Commerce Clause. And since Congress granted to federal courts exclusive jurisdiction over CERCLA cases, states raising such a bar effectively block private parties from recovering under CERCLA.[12] Thus, states benefit from CERCLA's enforcement power when it is directed at private parties, but states now are able to block contribution and cost-recovery actions initiated by private parties in cases in which the state has been found liable under CERCLA. Without the threat of a federal court action, states will be less likely to settle with private parties. More resources, will be spent litigating recovery actions and the cleanup of hazardous waste sites will be delayed.

This shift in power from the federal government to states is emblematic of the "new federalism" pursued in the other three recent federalism cases, *New York v. United States*, *United States v. Lopez*, and *City of Boerne v. Flores*.[13] Whereas the Court in *Lopez* and *New York* defined the contours of the Commerce Clause, the majority in *Seminole Tribe* used the clause to restrict federal jurisdiction. Justice John Paul Stevens, writing the first of two lengthy dissents in *Seminole Tribe*,[14] bluntly characterized the decision: "This case is about

---

7. U.S. CONST. art. I, § 8, cl. 3 ("Congress shall have Power . . . [t]o regulate Commerce with foreign Nations, and among the several States, and with the Indian Tribes.").

8. *Seminole Tribe*, 116 S. Ct. at 1131-32 ("Article I cannot be used to circumvent the constitutional limitations placed upon federal jurisdiction.").

9. Pennsylvania v. Union Gas Co., 491 U.S. 1 (1989), *overruled by Seminole Tribe*, 116 S. Ct. 1114. *See infra* Part I.B.1 and accompanying text.

10. *See infra* notes 162-63 and accompanying text.

11. U.S. CONST. amend. XI ("The Judicial power of the United States shall not be construed to extend to any suit in law or equity, commenced or prosecuted against one of the United States by Citizens of another State, or by Citizens or Subjects of any Foreign State.").

12. 42 U.S.C. § 9613(b) (1994).

13. As of the 1996-97 term, Chief Justice William Rehnquist and Justices Antonin Scalia, Sandra Day O'Connor, Andrew Kennedy, and Clarence Thomas comprise a majority bloc sympathetic to expanded states' rights on federalism issues. Justices John Paul Stevens, David Souter, Ruth Bader Ginsburg, and Stephen Breyer form a dissenting bloc on those issues.

14. *Seminole Tribe*, 116 S. Ct. at 1133, (Stevens, J., dissenting); *id.* at 1145 (Souter, J.,

power—the power of the Congress of the United States to create a private federal cause of action against a State . . . for the violation of a federal right."[15] He was emphatic about the impact of the decision, noting "it prevents Congress from providing a federal forum for a broad range of actions against States, from those sounding in copyright and patent law, to those concerning bankruptcy, environmental law, and the regulation of our vast national economy."[16]

CERCLA is illustrative of the impact of *Seminole Tribe* because it represents a vast, national expression of Congressional power directed at states and private parties. Already, at least one court has applied *Seminole Tribe's* restriction of jurisdiction in a CERCLA case in which a state was sued by a private party.[17] And as Justice Stevens predicted, the decision has reached into other areas of federal law.[18]

This Note analyzes the federalism issues in *Seminole Tribe* as they relate to CERCLA litigation and explores the new landscape many claimants may encounter. Part I surveys the federalism debate that formed the backdrop to *Seminole Tribe*. Part II briefly outlines the holdings and rationales of *Seminole Tribe's* majority and dissenting opinions. Part III analyzes the case's impact on CERCLA. Finally, Part IV suggests remedies and accommodations for restoring some measure of the balance lost in *Seminole Tribe* between state sovereignty and state accountability.

I. THE FEDERALISM CONTEXT

A. State Sovereign Immunity

Two key aspects of federalism coalesced in *Seminole Tribe*: state sovereign immunity and Congress' authority under the Commerce Clause. The Court's interpretations of these areas determine the reach of federal regulations such as CERCLA. The degree of sovereign immunity that states enjoy has been debated since before the Convention of 1787.[19] Likewise, the Court has rekindled the

---

dissenting).

15. *Id.* at 1133.

16. *Id.* at 1134.

17. Ninth Avenue Remedial Group v. Allis-Chalmers Corp., 962 F. Supp. 131 (N.D. Ind. 1997) (dismissing for lack of subject-matter jurisdiction pursuant to *Seminole Tribe* a CERCLA contribution action brought by private parties against an Indiana state agency).

18. *See infra* notes 108 and 109 and accompanying text.

19. *See* John J. Gibbons, *The Eleventh Amendment and State Sovereign Immunity: A Reinterpretation*, 83 COLUM. L. REV. 1889, 1897 (1983) (refuting the view that states enjoyed British-style sovereign immunity and did not surrender it in 1787, the author states "the relevant documents of the colonial period establish the absence of any expectation that governments were to be immune from suit."). For conflicting views on pre-Constitution views of state sovereign immunity, see THE FEDERALIST NO. 81, at 487-88 (Alexander Hamilton) (Rossiter 1961). Hamilton adds in *The Federalist* No. 32 that "as the plan of the Convention aims only at a partial Union or consolidation, the State Governments would clearly retain all the rights of sovereignty which they

debate on the limits of the Commerce Clause.[20]

While the Tenth Amendment,[21] which gives states residual powers, is the touchstone for most federalism issues, the controversy in *Seminole Tribe* centered on the Eleventh Amendment.[22] Textually, the amendment, passed in 1798, appears to divest federal courts of jurisdiction in diversity actions brought against a state by citizens of another state. Early Court opinions adopted this view.[23] Over the years, the Court has developed two other possible interpretations, (1) that it bars federal question jurisdiction in noncitizen suits against states,[24] or (2) that it blocks all private suits brought in federal court

---

before had, and which were not, by that act, *exclusively* delegated to the United States." THE FEDERALIST NO. 32, at 198 (Hamilton) (Rossiter 1961). Both sides of the contemporary federalism debate find currency with Hamilton's statements. *See Seminole Tribe*, 116 S. Ct. at 1130 (sovereign immunity referred to in THE FEDERALIST No. 81 applied to all federal jurisdiction over an unconsenting state); *id.* at 1166 (Souter, J., dissenting) (such immunity identified by Hamilton applied only to Citizen-State Diversity Clauses that would appear in Article III of the Constitution).

20. For thorough discussions on the "new federalism" attributed to the Rehnquist Court's trio of Commerce Clause cases highlighted in this Note (*see supra* notes 5 and 6), see Steven G. Calabresi, *A Government of Limited and Enumerated Powers: In Defense of* United States v. Lopez, 94 MICH. L. REV. 752 (1995); Lawrence Lessig, *Translating Federalism:* United States v. Lopez, 1995 SUP. CT. REV. 125; and Barry C. Toone & Bradley J. Wiskirchen, Note, *Great Expectations: The Illusion of Federalism After* United States v. Lopez, 22 J. LEGIS. 241 (1996).

21. U.S. CONST. amend. X ("The powers not delegated to the United States by the Constitution, nor prohibited by it to the States, are reserved to the States respectively, or to the people."). Much federalism jurisprudence focuses on the Tenth Amendment. For a modern example of the ongoing tension created by the amendment, see *U.S. Term Limits v. Thornton*, 115 S. Ct. 1842, 1855 (1995) ("[T]he Framers envisioned a uniform national system, rejecting the notion that the Nation was a collection of States, and instead creating a direct link between the National Government and the people of the United States."); *id.* at 1875 (Kennedy, J., concurring) ("[T]here exists a federal right of citizenship . . . with which the States may not interfere."); but compare, *id.* at 1875 (Thomas, J., dissenting) (In characterizing the "reserved" powers of states, "[t]he ultimate source of the Constitution's authority is the consent of the people of each individual State, not the consent of the undifferentiated people of the Nation as a whole.").

22. *See supra* note 11 and accompanying text.

23. *See* Cohens v. Virginia, 19 U.S. 264, 382 (1821) (Chief Justice Marshall writing for the Court concluded "a case arising under the constitution or laws of the United States, is cognizable in the Courts of the Union, whoever may be the parties to that case."); Osbourn v. Bank of United States, 22 U.S. 738, 847 (1824) ("The eleventh amendment of the constitution has exempted a State from the suits of citizens of other States." The opinion makes no reference to federal question jurisdiction.).

24. The Court has never held to this interpretation, but it has appeared as dicta in opinions. *See* Atascadero State Hospital v. Scanlon, 473 U.S. 234, 288 n.41 (1985) (Brennan, J., dissenting) ("When the Court is prepared to embark on a defensible interpretation of the Eleventh Amendment consistent with its history and purposes, the question whether the Amendment bars federal-question or admiralty suits by a noncitizen or alien against a State would be open.").

against unconsenting states regardless of citizenship.[25]

In *Seminole Tribe*, the Court narrowly affirmed its broadest reading of the Eleventh Amendment, a view articulated more than a century earlier in *Hans v. Louisiana*.[26] The *Hans* Court ruled that the grant of immunity extended to all federal suits brought against states by private parties.[27] The Court reasoned that, since a state could invoke sovereign immunity against a noncitizen suing under federal question jurisdiction, it enjoyed the same immunity in such suits involving its own citizens.[28] This view is the law today, but with the sizable exception that states may not use the Eleventh Amendment to bar private suits authorized by Congress under the Fourteenth Amendment.[29] Eighteen years after *Hans*, in *Ex parte Young*,[30] another case central to *Seminole Tribe*, the Court shifted the federalism equation in the federal government's favor by providing private parties with injunctive relief against state officers who violate federal law in cases in which citizens may not sue the state directly.[31]

B. Setting the Stage for Seminole Tribe

Against the backdrop of the seemingly unchecked expansion of Congressional power under the Commerce Clause that continued into the 1980s, the Court bitterly separated into two factions in the name of federalism. In four

---

25. Hans v. Louisiana, 134 U.S. 1, 10 (1890). In *Hans*, Louisiana sought to repudiate Reconstruction bonds held by one of its citizens and issued by the state during federal occupation. *See also Seminole Tribe*, 116 S. Ct. at 1131 ("Even when the Constitution vests in Congress complete lawmaking authority over a particular area, the Eleventh Amendment prevents congressional authorization of suits by private parties against unconsenting States.").

26. 134 U.S. 1. Justice Souter, dissenting in *Seminole Tribe*, argued that the *Hans* decision was a concession to the emerging Southern state governments following the end of Reconstruction. The Court in *Hans*, Souter argued, found a way to allow Louisiana to bar the suit rather than to uphold existing Eleventh Amendment doctrine and see its decision go unenforced after federal troops had left the South. 116 S. Ct. at 1155 (Souter, J., dissenting).

27. *Hans*, 134 U.S. at 10.

28. *Id.*

29. Fitzpatrick v. Bitzer, 427 U.S. 445, 455 (1976) (award of retirement benefits to retired state employees found to have been discriminated against on the basis of sex under the state's retirement plan, in violation of a provision of the Civil Rights Act of 1964, *as amended*, 42 U.S.C. § 2000(e) (1994), was not barred by the Eleventh Amendment; section 5 of the Fourteenth Amendment grants Congress authority to enforce "by appropriate legislation" limitations on state authority).

30. 209 U.S. 123 (1908).

31. *Id.* at 159-60 (The federal government can enjoin a state official when the officer's action "is simply an illegal act upon the part of a state official in attempting, by the use of the name of the state, to enforce a legislative enactment which is void because unconstitutional . . . . The state has no power to impart to him any immunity from responsibility to the supreme authority of the United States.").

major Commerce Clause cases,[32] the factions remained largely intact and are so today. Chief Justice Rehnquist and Justices Kennedy, O'Connor, Scalia, and Thomas, with slight variance,[33] formed one side, which was successful in limiting Congress' Commerce Clause power in a manner not seen in almost sixty years. Justice Stevens, the lone member of the other faction to appear in all four decisions, aligned with Justices Souter, Breyer, and Ginsburg to form the other faction. Retired Justices Brennan, Marshall, Blackmun, and White generally supported the latter group.

1. Pennsylvania v. Union Gas Co.—To reach its decision in *Seminole Tribe*, the Court expressly overruled *Union Gas*.[34] Many of the constitutional arguments raised in *Union Gas* were repeated in *Seminole Tribe*. *Union Gas* was a case of first impression for the Court. The issue was whether Congress had authority under the Commerce Clause to subject an unconsenting state to suit in federal court.

The United States brought a CERCLA action against the operator of a coal gasification plant to recover cleanup costs.[35] The operator initiated a third-party suit against Pennsylvania. The district court dismissed the operator's suit,[36] accepting the state's claim that the Eleventh Amendment barred the suit. The Third Circuit affirmed the district court's decision, holding it found no clear Congressional intent to make states liable for monetary damages in CERCLA claims.[37] While the Supreme Court considered the gas company's certiorari petition, Congress amended CERCLA, including a contribution action as part of the Superfund Amendment and Reauthorization Act of 1986.[38] The high court granted certiorari, vacated the Third Circuit opinion, and remanded it to the Court of Appeals for consideration of the amendments.[39] The Third Circuit then found the amendments showed clear intent to subject states to damages under CERCLA.[40] The case then went back to the Supreme Court.[41]

The Court embarked on a two-step analysis that it would later repeat in *Seminole Tribe*. First, it looked at whether Congress had made its intent to abrogate state sovereign immunity "unmistakably clear."[42] Second, the Court decided whether abrogation under the Commerce Clause of state sovereign

---

32. Pennsylvania v. Union Gas Co., 491 U.S. 1 (1989), *overruled by Seminole Tribe*, 116 S. Ct. 1114; *New York*, 505 U.S. 144; *Lopez*, 514 U.S. 549; *Seminole Tribe*, 116 S. Ct. 1114.

33. Souter joined the majority in *New York*, 505 U.S. 144.

34. *Seminole Tribe*, 116 S. Ct. at 1131.

35. 42 U.S.C. §§ 9604 & 9606 (1994).

36. United States v. Union Gas Co., 575 F. Supp. 949 (1983).

37. United States v. Union Gas Co., 792 F.2d 372 (1986).

38. Pub.L. 99-499, 100 Stat. 1613 (1986), *now codified as* 42 U.S.C. § 9613(f).

39. 479 U.S. 1025 (1987).

40. 832 F.2d 1343 (1986).

41. 485 U.S. 958 (1988).

42. *See Atascadero*, 473 U.S. at 242 ("Congress may abrogate the States' constitutionally secured immunity from suit in federal court only by making its intention unmistakably clear in the language of the statute.").

immunity is constitutionally valid. A plurality consisting of Justices Brennan, Marshall, Blackmun, and Stevens held that the plain language of CERCLA indicated that states fell within the definitions of "person" and "owner or operator,"[43] all of whom were subject to liability for remedial costs. Also, the Court noted a provision exculpating state and local governments for actions taken in emergencies. The provision reads "[T]his paragraph shall not preclude liability for costs or damages as a result of gross negligence or intentional misconduct."[44] The Court then concluded that Congress, in enacting CERCLA, used valid power under the Commerce Clause to subject Pennsylvania to suit, overriding the state's claim of sovereign immunity under the Eleventh Amendment.[45] Justice Brennan, who wrote for the Court, reasoned that the states effectively waived their immunity when they "granted Congress the power to regulate commerce;"[46] that Congress' plenary authority to regulate interstate commerce, analogous to its Fourteenth Amendment enforcement authority, allowed it to abrogate state sovereign immunity;[47] and that Congress' vast power under the Commerce Clause displaced state authority and sometimes precludes state regulation, even in areas in which the federal government has chosen not to act.[48]

Justice Brennan feared the impact of denying private claims against states.[49] Such a reading would frustrate Congress' "legitimate objectives under the Commerce Clause."[50] In applying this reasoning to CERCLA, Brennan wrote, "the case before us brilliantly illuminates these points. The general problem of environmental harm is often not susceptible of a local solution."[51] He compared *Union Gas* to *Philadelphia v. New Jersey,*[52] a 1978 case in which the Court held that a New Jersey statute regulating out-of-state solid waste and exempting in-state waste violated the Commerce Clause.[53] Brennan argued the proposition that some environmental problems can only be solved by the federal government,

---

43. 42 U.S.C. § 9601(21) (1994) (persons definition includes "state"); *id.* § 9601(20)(A) (owner or operator); *id.* §9607(a) (who is liable).

44. *Id.* § 9607(d)(2).

45. *Union Gas*, 491 U.S. at 23. Justice White concurred in the judgment on the constitutional question and dissented on the intent to abrogate issue. *Id.* at 45-56.

46. *Id.* at 14. *See* Parden v. Terminal Railway of Alabama Docks Dept., 377 U.S. 184, 192 (1964) ("By empowering Congress to regulate commerce, then, the States necessarily surrendered any portion of their sovereignty that would stand in the way of such regulation."), *overruled on other grounds*, Welch v. Texas Dept. of Highways and Public Transp., 483 U.S. 468 (1987).

47. *Union Gas*, 491 U.S. at 15.

48. *Id.* at 20.

49. *Id.* ("[A] conclusion that Congress may not create a cause of action for money damages against the States would mean that no one could do so.").

50. *Id.* This is precisely the dilemma faced by potential plaintiffs in third-party cases post-*Seminole Tribe*. *See infra* Part III.D.

51. *Id.* at 20

52. 437 U.S. 617 (1978).

53. *Union Gas*, 491 U.S. at 21 (citing Philadelphia v. New Jersey, 437 U.S. at 626-29).

adding "often those solutions, to be satisfactory, must include a cause of action for money damages.[54] CERCLA was a comprehensive remedy to a national problem of hazardous waste that earlier efforts failed to solve, he argued.[55] Congress encouraged voluntary cleanup efforts and induced such action authorizing private parties to recover cleanup costs from other potentially responsible parties.[56] To exclude states from the scheme, he argued, would thwart the purpose of CERCLA:

> If States, which comprise a significant class of owners and operators of hazardous-waste sites, need not pay for the costs of cleanup, the overall effect on voluntary cleanups will be substantial. This case thus shows why the space carved out for federal legislation under the commerce power must include the power to hold States financially accountable not only to the Federal Government, but to private citizens as well.[57]

The counterweight to the Court's decision was Justice Scalia's dissent.[58] The opinion attracted four-fifths of the majority that would, with the addition of Justice Thomas, decide *Seminole Tribe* seven years later. Justice Scalia based his opinion on a broad reading of *Hans v. Louisiana,*[59] in which the Court held a state could invoke sovereign immunity in all federal suits brought by private parties.[60] The dissent set out a constitutional analysis that would be repeated in *Seminole Tribe*. Scalia called *Hans* a "landmark case" that repudiated the "comprehensive" reading of the Eleventh Amendment, a view that limits state sovereign immunity to citizen-state diversity actions.[61] Scalia argued the Eleventh Amendment stands for more than the textual grant of sovereign immunity in some diversity cases—it shows that states enjoyed substantial sovereign immunity before the Convention of 1787 and that immunity was not entirely eliminated by Article III of the Constitution.[62]

Scalia's characterization of *Hans* in *Union Gas* presages the ruling in

---

54. *Id.*

55. *Id.*

56. *Id.* at 21-22.

57. *Id.* at 22.

58. *Id.* at 29-45 (Scalia, J., dissenting). The case featured three other opinions, *see id.* at 23-29 (Stevens, J., concurring) (Stevens argued that Congress had properly balanced environmental protections with subjecting states to damages actions, that the "judicial power" of the United States extends to such suits, and—that even if the Court had disagreed with the balance Congress struck—the Court may not disregard Congress' "express decision to subject the States to liability under federal law."); *id.* at 45-56 (White, J., concurring) (White concurred in the constitutional issue without comment but disagreed with the majority's assertion that Congress had made its intent "unmistakably" clear.); *id.* at 57 (O'Connor, J., dissenting) (no substantive comments.).

59. 134 U.S. 1 (1890).

60. *Id.* at 10.

61. *Union Gas*, 491 U.S. at 31.

62. *Id.* at 31-32.

Seminole Tribe.[63] Conversely, Justice Stevens, in his *Union Gas* concurrence, rebukes *Hans* as having created a second Eleventh Amendment, a "judicially created doctrine of state sovereign immunity."[64] Both he and Justice Souter, in their *Seminole Tribe* dissents,[65] attacked the continued reliance on *Hans*.

2. New York v. United States.—This case is the first of the four major "new federalism" decisions.[66] Commentators have argued that these cases, *New York*, *Lopez*, *Seminole Tribe*, and *City of Boerne*, all upholding state power at the expense of Congress, could become the legacy of the Rehnquist Court.[67] In *New York*, the Court, for just the second time since 1936, invalidated a federal law on Tenth Amendment grounds.[68] Writing for the majority, Justice O'Connor said the Court faced "one of our Nation's newest problems of public policy and perhaps our oldest question of constitutional law."[69] The new problem was regulation of low-level radioactive waste; the old constitutional issue was federalism. New York brought suit against the federal government, challenging provisions of the Low-Level Radioactive Waste Policy Act of 1985.[70] The act offered three sets of incentives to induce state compliance. The first rewarded states monetarily for accepting out-of-state radioactive waste.[71] The second allowed states to increase the cost of waste disposal and cut off-site access for waste generated in states not participating in the federal program.[72] The third imposed a Faustian choice on states: take title of waste generated in their borders

---

63. *Seminole Tribe*, 116 S. Ct. at 1127 ("The [*Union Gas*] plurality's rationale also deviated sharply from our established federalism jurisprudence and essentially eviscerated our decision in *Hans*.").

64. *Union Gas*, 491 U.S. at 23.

65. *Seminole Tribe*, 116 S. Ct. at 1137, 1144.

66. 505 U.S. 144. The other two are *Seminole Tribe*, 116 S. Ct. 1114, and *Lopez*, 514 U.S. 549.

67. *See, e.g.*, Loren Singer, *Scholar Calls Supreme Court's Clarity on Some Issues, Indecision on Others the "Most Remarkable Thing" About the 1995-96 Term*, Oct. 22, 1996, WEST LEGAL NEWS 11214 (covers remarks made Oct. 19, 1996, by Professor Charles H. Whitebread at the University of Minnesota Law School: although *Seminole Tribe* was the "least important" case in the "new federalism" trend, "[t]hat trend will be Chief Justice Rehnquist's 'major legacy' if the decisions are followed in the future."); Linda Greenhouse, Analysis, *Rehnquist Legacy: States Rights*, ORANGE COUNTY (Calif.) REGISTER, April 14, 1996, at A25. ("When the Rehnquist Court passes into history . . . a reshaping of the federal-state balance may prove his most enduring legacy."); *Reining in Congress*, WASH. POST, June 28, 1997 ("In recent years, the Rehnquist Court has curtailed the power of Congress.").

68. The other Tenth Amendment case was *National League of Cities v. Usery*, 426 U.S. 833 (1976), which was expressly overruled by *Garcia v. San Antonio Metropolitan Transit Authority*, 469 U.S. 528, 546-47 (1985) (rejecting the notion in *Usery* that constitutionality of a federal regulation depends on whether courts find it outside of "integral" or "traditional" state authority).

69. *New York*, 505 U.S. at 149.

70. Pub.L. 99-240, 99 Stat. 1842 (1985), *now appears as* 42 U.S.C. §§ 2021b-2021j (1994).

71. 42 U.S.C. § 2021e(d)(2)(A).

72. *Id.* § 2021e(e)(2)(A-D).

or regulate according to Congress' scheme.[73]

The Court found that the "monetary" incentives were a valid exercise of Congress' power under the Spending Clause[74] and the Commerce Clause;[75] that the "access" incentives were examples of "cooperative federalism"—a conditional exercise of Commerce Clause power; but that the "take title" provisions violated the Tenth Amendment.[76] States had no choice but to assent to the regulations, an affront to their sovereignty, O'Connor wrote.[77] The government argued unsuccessfully that the scheme was valid under the Commerce Clause because it merely sought to arbitrate interstate conflicts in the disposal of radioactive waste. Justice White, in a dissent, criticized the "formalistically rigid obeisance to 'federalism' " employed by the majority.[78] He called it "the ultimate irony" that the Tenth Amendment, an assertion of states' rights, was employed to invalidate a scheme aimed at state and local solutions to a national environmental problem.[79] Justice Stevens was more direct in his dissenting assault. Federalism aside, the national government clearly had the power, in his opinion, to order state compliance in a host of areas of national concern, including radioactive waste management:[80] "[t]he notion that Congress does not have the power to issue a simple command to state governments to implement legislation enacted by Congress is incorrect and unsound."[81]

The impact of *New York* on Congress' Commerce Clause power was subtle. However, the expansive reading of the Tenth Amendment, which seems to have slightly upset the regulatory scheme in *New York*, presaged a direct assault on Congress' power three years later.

3. United States v. Lopez.—In a landmark case, the Court, for the first time since 1936, held a federal law unconstitutional because it exceeded Congress' authority under the Commerce Clause.[82] *Lopez* brought the "new federalism" debate to the forefront[83] and provided the backdrop for *Seminole Tribe*. After nearly sixty years of unchecked Congressional action under the Commerce Clause, the Court, led by Chief Justice Rehnquist, stoked the federalism debate

---

73. *Id.* § 2021e(d)(2)(C).

74. *New York*, 505 U.S. at 173 (interpreting U.S. CONST. art. I § 8, cl. 1, "Congress shall have Power to lay and collect Taxes, Duties, Imposts and Excises, to pay the Debts and provide for the common Defence and general Welfare of the United States; but all Duties, Imposts and Excises shall be uniform throughout the United States.").

75. *Id.* at 174 (interpreting U.S. CONST. art. I, § 8, cl. 3.); *see supra* note 7 and accompanying text.

76. *Id.* at 177 (interpreting U.S. CONST. amend X); *see supra* note 21 and accompanying text.

77. *Id.*

78. *Id.* at 210.

79. *Id.*

80. *Id.* at 211.

81. *Id.*

82. 514 U.S. 549 (1995).

83. *See supra* note 20 and accompanying text.

and revisited two centuries of history.[84]

In *Lopez*, a defendant was convicted of possessing a firearm in a school zone in violation of the Gun-Free School Zones Act of 1990.[85] The Fifth Circuit reversed the conviction, finding the statute beyond the power under the Commerce Clause.[86] The Supreme Court affirmed the decision, holding the statute was criminal in nature and "ha[d] nothing to do with 'commerce' or any sort of economic enterprise."[87] It also held that the statute contained no jurisdictional element, "which would ensure, through case-by-case inquiry" the nexus to commerce.[88]

The opinion, like those in *New York* and *Seminole Tribe*, featured lengthy explorations of constitutional history and precedent. "We start with first principles," Rehnquist wrote,[89] launching into two centuries of federalism analysis. The Court based its opinion on existing Commerce Clause doctrines but read into those doctrines a more active role for itself in deciding the validity of congressional action. Citing precedent for each, the Court gave three categories in which Congress may regulate commerce: (1) to protect the channels for interstate commerce,[90] (2) to protect the instrumentalities of commerce,[91] and (3) to control activities that have a "substantial relation" to interstate commerce.[92]

The Court analyzed the firearms statute under the third category and found it outside of the Commerce Clause's grant of authority.[93] Implicit in the Court's

---

84. Joining Rehnquist were Justices Scalia, Kennedy, O'Connor, and Thomas, the same majority in *Seminole Tribe*. Kennedy and Thomas wrote concurring opinions. Dissenting were Justice Breyer (joined by Stevens, Souter, and Ginsburg), and Justices Stevens and Souter in separate opinions. The dissenters were the same in *Seminole Tribe*.

85. 18 U.S.C. § 922(q) (1994), *amended by* Pub. L. 103-322 (1994).

86. 2 F.3d 1342, 1367, 1368 (5th Cir. 1993).

87. *Lopez*, 514 U.S. at 561.

88. *Id.*

89. *Id.* at 552.

90. *See* United States v. Darby, 312 U.S. 100, 113 (1941) ("While manufacture is not of itself interstate commerce the shipment of manufactured goods interstate is such commerce and the prohibition of such shipment by Congress is indubitably a regulation of the commerce.").

91. *See* Houston, E. & W.T.R. Co. v. United States, 234 U.S. 342, 351 (1914) (Shreveport Rail Rate Cases) (Congress' authority "extending to these interstate carriers as instruments of interstate commerce, necessarily embraces the right to control their operations in all matters having such a close and substantial relation to interstate traffic that the control is essential or appropriate to the security of that traffic, to the efficiency of the interstate service, and to the maintenance of conditions under which interstate commerce may be conducted upon fair terms and without molestation or hindrance.").

92. *See* NLRB v. Jones & Laughlin Steel Corp., 301 U.S. 1, 37 (1937) ("Although activities may be intrastate in character when separately considered, if they have such a close and substantial relation to interstate commerce that their control is essential or appropriate to protect that commerce from burdens and obstructions, Congress cannot be denied the power to exercise that control.").

93. *Lopez*, 514 U.S. at 561.

opinion is that the statute also ran afoul of the Constitution because it sought to regulate non-economic activity.[94] In the chief dissent, Justice Breyer argues that the legislation is sustained by the Commerce Clause because Congress had a "rational basis" for concluding that a connection existed between gun possession near schools and interstate commerce.[95]

The opinion that focuses most acutely on the elusive federalism balance, however, was Justice Kennedy's concurrence. While joining in the "necessary though limited holding"[96] of the majority, Kennedy commented that federalism issues are murkier than those arising under the other constitutional pillars: separation of powers, checks and balances, and judicial review.[97] In the Court's attempt to strike the proper balance, Kennedy wrote, "[o]ur ability to preserve this principle under the Commerce Clause has presented a much greater challenge" compared to the Court's role in other doctrines.[98] In the end, Kennedy agrees that the statute should fail for its lack of a commercial nexus, that it upsets the constitutional balance.[99] But he is pointedly cautious in observing that "[w]hile the intrusion on state sovereignty may not be as severe in this instance as in some of our recent Tenth Amendment cases, the intrusion is nonetheless significant."[100] Commentators have noted that Kennedy and Justice O'Connor hold the balance of power on the Court.[101] In future federalism cases, it is possible that the cautious view articulated in *Lopez* by Kennedy will align with the "liberal bloc," the *Lopez* dissenters, when the balance tips too far away from the federal government.[102]

If *New York v. United States* raised any question about the depth of the majority's resolve in shifting the federalism balance toward the states, *Lopez* provided a clear answer. By reining in Congress on the Commerce Clause for the first time in almost six decades, the Court opened the door for an array of challenges to federal authority claimed in the name of interstate commerce.[103] Understandably, the ruling has caused confusion in the district courts. One district court in Alabama relied on *Lopez* to invalidate retroactive provisions of CERCLA because the regulated activity—hazardous waste deposited by prior

---

94. *Id.*
95. *Id.* at 631.
96. *Id.* at 568.
97. *Id.* at 575.
98. *Id.* at 579.
99. *Id.* at 583.
100. *Id.*
101. *See* Singer, *supra* note 67 and accompanying text.
102. *Id.*
103. For example, a federal district court, in *Brzonkala v. Virginia Polytechnic & State University*, 935 F. Supp. 779, 787 (W.D. Va. 1996), used the two-step analysis in *Lopez* to strike down The Violence Against Women Act, 42 U.S.C. § 13981 (1994), finding the legislation an invalid use of Commerce Clause powers because the activity was non-economic and the statute lacked a jurisdictional element that would ensure case-by-case inquiry of its application to properly regulated activity.

users—did not bear a substantial relation to commerce and because CERCLA lacked a jurisdictional element that would ensure case-by-case inquiry of its relevance to commerce.[104] However, an Illinois district court took the opposite approach, finding CERCLA was a valid exercise of Commerce Clause authority because Congress had a rational basis for determining that hazardous waste activity "substantially affects" interstate commerce.[105] Thus, *Lopez* put federal courts in the position of applying its Commerce Clause test to different factual situations on a case-by-case basis. By contrast, the Supreme Court's next significant federalism case would provide clearer guidance to further its ends.

II. SEMINOLE TRIBE OF FLORIDA V. FLORIDA

Just as the scope of *Lopez* is not limited to school safety, *Seminole Tribe's* significance goes far beyond the specific controversy that brought it before the Court: the regulation of Indian gaming rights.[106] In one sense, this third expression of the "new federalism" fits into a trend from general to specific established by the two earlier decisions, *Lopez* and *New York*. In *New York*, the Court outlined the context for restricting Commerce Clause power, and, in *Lopez*, the Court struck down a congressional provision in that context. In *Seminole Tribe*, the Court cuts off Commerce Clause power at the source with a broad reading of the Eleventh Amendment. In another sense, the case can be seen as an intersection of two federalism issues—the Commerce Clause and state sovereign immunity—that had simmered below the surface for years.

Specifically, the case changed the rules for the regulation of Indian gaming activities and overruled a key CERCLA case.[107] Lower court cases in both arenas were immediately affected.[108] Like *Lopez*, the decision set off a cavalcade of

---

104. United States v. Olin, 927 F. Supp. 1502, 1532-33 (S.D. Ala. 1996), *rev'd*, 107 F.3d 1506 (11th Cir. 1997). This case was roundly criticized in by Richard Lazarus in *Striking Down Retroactive Liability*, ENVIRONMENTAL FORUM, July 4, 1996, at 8.

105. United States v. N.L. Industries, 936 F. Supp. 545, 563 (S.D. Ill. 1996).

106. While *Seminole Tribe* did not dominate the front pages, most major U.S. daily newspapers seized upon the federalism implications in coverage and commentary. *See, e.g.*, Editorial, *Seminoles and State Sovereignty*, WASH. POST, March 30, 1996, at A16 ("[I]t is a broad victory for those who believe the federal government has been encroaching on the prerogatives of the states in a manner never contemplated by the Founders."); Editorial, *Another Judicial Victory for Authority of the States*, L.A. TIMES, March 29, 1996, at 8 ("Congress' power to address problems of such obvious federal interest as violence at abortion clinics, narcotics, 'deadbeat dads,' hazardous waste dumps or pistols in the hands of ex-felons may be in question."); Editorial, *Restoring Federalism*, DETROIT NEWS, March 29, 1996 at A8 ("Now that the Eleventh Amendment has been rediscovered, we hope the rediscovery of other sections of the Constitution . . . such as the 10th Amendment, won't be far behind.").

107. *Seminole Tribe*, 116 S. Ct. 1114, 1128 (overruling *Union Gas*, 491 U.S. 1).

108. Two district courts divested themselves of jurisdiction in a private CERCLA action against a state. *See Ninth Avenue*, 962 F. Supp. at 131; Prisco v. New York, 1996 WL 596546 (S.D.N.Y. Oct. 16, 1996). For cases in which district court rulings dismissing complaints based

decisions covering several other sectors of federal law in the months following its announcement. In most cases, the basic holding of *Seminole Tribe*—that Congress may not use Article I powers to abrogate state sovereign immunity from federal suits brought by private parties—was used to convince federal courts to divest themselves of jurisdiction.[109]

---

on state sovereign immunity claims pursuant to the Indian Gaming Regulatory Act, 25 U.S.C. § 2710 (1994) were affirmed following *Seminole Tribe*, see Spokane Tribe of Indians v. Washington, 790 F. Supp. 1057 (E.D. Wash. 1991), *rev'd*, 28 F.3d 991 (9th Cir. 1994), *vacated and remanded*, 116 S. Ct. 1410 (1996), *aff'd*, 91 F.3d 1350, 1351 (9th Cir. 1996); Ponca Tribe of Oklahoma v. Oklahoma, 834 F. Supp. 1341 (W.D. Okla. 1992), *aff'd in part, rev'd in part*, 37 F.3d 1422 (10th Cir. 1994), *vacated and remanded*, 116 S. Ct. 1410 (1996), *aff'd*, 89 F.3d 690 (10th Cir. 1996).

109. As Justice Stevens predicted in *Seminole Tribe*, 116 S. Ct. at 1134, the ruling has had a dramatic effect on several areas of federal law. For the impact on copyright cases, see *Genentech, Inc. v. Regents of the University of California*, 939 F. Supp. 639, 642 (S.D. Ind. 1996) (district court in a declaratory judgment action found that actions brought pursuant to the Patents and Copyright Clause, U.S. CONST. art. I, § 8, cl. 8, against state defendants were subject to state sovereign immunity claims; finding no waiver of immunity, the court granted dismissal in favor of a state university); Chavez v. Arte Publico Press, 59 F.3d 539 (5th Cir. 1995) (author brought a copyright infringement case, alleging a state university violated provisions of the Lanham Act, 15 U.S.C. § 1122 (1994) and the Copyright Act, 17 U.S.C. § 511(a) (1994); Fifth Circuit denied Texas' claim of sovereign immunity holding that the Patents and Copyright Clause gave Congress the authority to abrogate state sovereign immunity. *Chavez*, 59 F.3d at 546, but after *Seminole Tribe*, the Supreme Court granted certiorari, 116 S. Ct. 1667 (1996), and vacated and remanded the decision.). For an excellent discussion of *Chavez* and the impact *Seminole Tribe* may have on state universities, see Douglas Lederman, *Supreme Court Gives Public Universities New Protection Against Lawsuits*, CHRONICLE OF HIGHER EDUCATION, November 8, 1996, at A33.

For *Seminole Tribe's* impact on state-employee cases, see *Close v. New York*, No. 94-CV-0906, 1996 WL 481550 (N.D.N.Y. Aug. 19, 1996) (following *Seminole Tribe's* ban on abrogation of state sovereign immunity under the Commerce Clause, a federal court divested itself of jurisdiction in a case brought by state employees against New York under provisions of the Fair Labor Standards Act, 29 U.S.C. §§ 201-219 (1994).); Wilson-Jones v. Caviness, 99 F.3d 203, 206 (6th Cir. 1996) (case brought by Ohio state employees dismissed; relying on *Seminole Tribe*, court found the claim arose under the Commerce Clause, not the Equal Protection Clause, U.S. CONST. amend. XIV § 1, and therefore the court did not have jurisdiction).

For the impact of cases brought under the Bankruptcy Clause, U.S. CONST. art I, § 8, cl. 4 (Congress has the power "[t]o establish . . . uniform Laws on the subject of Bankruptcies throughout the United States.), see York-Hanover Developments, Inc., v. Florida Dept. of Revenue, 201 B.R. 137, 141 (Bankr. E.D.N.C. 1996) (Chapter 7 trustee sought the return of alleged fraudulent transfers from the Florida Department of Revenue. The court found that Congress could not, as an exercise of Bankruptcy Clause power, abrogate state sovereign immunity under 11 U.S.C. § 106(a) (1994).).

See also Gorka v. Sullivan, 82 F.3d 772, 775 (7th Cir. 1996) (Medicaid recipients brought state and federal claims against Indiana in state court. Indiana removed case to federal court and claimed sovereign immunity to some of the claims. The Seventh Circuit said that although *Seminole Tribe* broadened state sovereign immunity, states may not use it as a sword and a shield. States may

A. The Case in General

The dispute arose from the attempts by the Seminole Tribe of Indians to establish commercial gaming activities in Florida pursuant to the Indian Gaming Regulatory Act (IGRA),[110] which requires tribes to enter into a valid compact with the state in which the activities will be located.[111] The states have a duty to negotiate in good faith toward the formation of such a pact.[112] The act authorizes tribes to sue states in federal court to compel states to perform their duty.[113] Congress passed the act pursuant to the Indian Commerce Clause, which is found in the same sentence that authorizes Congressional regulation of commerce with foreign nations and "among the several states."[114]

The Seminoles brought suit in federal district court, alleging the state failed to negotiate in good faith. The state moved to dismiss the complaint, arguing the suit violated the state's sovereign immunity under the Eleventh Amendment. The district court denied the state's motion.[115] The Eleventh Circuit reversed the district court decision, recognizing the state's Eleventh Amendment bar and remanded the case with orders to dismiss.[116] The circuit court also ruled the tribe could not force the state to negotiate under the *Ex parte Young* doctrine.[117] The Supreme Court, in a 5-4 decision, affirmed the Eleventh Circuit's order to dismiss the case.[118] In reaching its decision, the Court expressly overruled *Pennsylvania v. Union Gas,* [119] in which the Court had held that Congress has Commerce Clause authority to subject unconsenting states to suits brought by private parties in federal court.[120]

B. The Majority's Analysis

In many ways, *Seminole Tribe* can be read as a continuation of the debate

---

not remove a case to a federal forum and then contend that no relief may be granted when it gets there.).

110. 25 U.S.C. § 2710 (1994).

111. *Id.* § 2710(d)(1)(C).

112. *Id.* § 2710(d)(3)(A).

113. *Id.* § 2710(d)(7).

114. U.S. CONST. art. I, § 8, cl. 3; *see supra* note 7 and accompanying text. The Court accepts the view in *Union Gas,* 491 U.S. 1, that no "principled distinction" exists between the effect of the Indian Commerce Clause and the Interstate Commerce Clause, but it notes that, if anything, the Indian Commerce Clause, which the Court construes in Florida's favor, represents a greater grant of sovereignty to the federal government than the Interstate Commerce Clause. *Seminole Tribe*, 116 S. Ct. at 1126, 1127.

115. 801 F. Supp. 655 (S.D. Fla. 1992).

116. 11 F.3d 1016 (11th Cir. 1994).

117. *Id.* at 1028, 1029. *See also Ex parte Young*, 209 U.S. at 159-60.

118. *Seminole Tribe*, 116 S. Ct. at 1122.

119. *Id.* at 1128.

120. *Union Gas*, 491 U.S. at 23.

begun in *Union Gas*, only the sides have switched position and the players have changed. With Justice Thomas joining the Court and the retirement of Justice Brennan, who wrote the *Union Gas* opinion, the majority was free to dispense with *Union Gas* as an aberration. As Thomas and Justices Kennedy, O'Connor, and Scalia joined him, Chief Justice Rehnquist penned a spare opinion compared to the tomes featured in the other recent federalism cases. For the first holding, that the IGRA-grounded suit was an invalid exercise of Commerce Clause authority, Rehnquist pursued a two-step analysis: did Congress "unmistakably" intend to abrogate state sovereign immunity and was it acting under valid constitutional authority? Rehnquist answered the former question in the affirmative, as he found express terms in the IGRA that subjected the state to suit.[121] Rehnquist then addressed the second inquiry—whether the abrogation was valid under the Commerce Clause.

As in *Union Gas*, the logic of the majority and dissenting opinions flowed from how each side characterized *Hans*,[122] in which the Court held that states enjoy immunity from all federal suits brought by private parties.[123] The majority endorsed the rule in *Hans* "essentially eviscerated" in *Union Gas*.[124] But, in *Seminole Tribe*, the majority criticized *Union Gas* because it "deviated sharply from our established federalism jurisprudence;" as a plurality opinion, its rationale was not sustained by a majority of the Court;[125] and it was a "solitary departure from established law."[126] In defending *Hans*, the majority argued that the *Hans* Court had a greater vantage point from which to infer the nature of pre-Eleventh Amendment state sovereign immunity than the dissent, and therefore *Hans'* extension of the amendment to federal question jurisdiction was a more valid point of reference than *Chisholm v. Georgia*,[127] the case cited for the same purpose by the dissent.[128] Specifically, the majority accused the dissent of

---

121. *Seminole Tribe*, 116 S. Ct. at 1124. ("[W]e think that the numerous references to the "State" in the text of [25 U.S.C. § 2710(d)(7)(B) (1994)] make it indubitable that Congress intended through the Act to abrogate the States' sovereign immunity from suit." For example, the court cited § 2710(d)(7)(B)(ii)(II), which "provides that if a suing tribe meets its burden of proof, then the burden of proof shall fall upon the State." 116 S. Ct. at 1124.).

122. 134 U.S. 1, 10 (1890), *see supra* note 25 and accompanying text.

123. *Id.*

124. *Seminole Tribe*, 116 S. Ct. at 1127. In *Union Gas*, the Court upheld a third-party claim against Pennsylvania in a CERCLA action. 491 U.S. at 24.

125. *Seminole Tribe*, 116 S. Ct. at 1127.

126. *Id.* at 1128. *See also* a case cited by the Court, *Schweiker v. Chilicky*, 487 U.S. 412, 423 (1988) ("When the design of a Government program suggests that Congress has provided what it considers adequate remedial mechanisms for constitutional violations that may occur in the course of its administration, we have not created additional . . . remedies.").

127. 2 U.S. (2 Dall.) 419, 429, 448 (1793) (Georgia could not invoke sovereign immunity in an assumpsit case brought by a South Carolina citizen seeking repayment of Revolutionary War loans.).

128. *Seminole Tribe*, 116 S. Ct. at 1130.

putting forward a new theory of state sovereign immunity,[129] a charge the dissent returned almost verbatim.[130]

For its second holding, the majority relied on the text of IGRA and recent precedent in holding that injunctive relief under the *Ex parte Young* doctrine was not available to the Seminoles.[131] The doctrine allows relief against a state officer found to be in violation of federal law, but it is limited to situations in which Congress has not prescribed a detailed remedial scheme.[132] Such an "intricate" procedure for remedy existed in IGRA, the Court found.[133] This holding left the Seminoles with a federal right and no judicial forum in which to prosecute it, a position in which many claimants now find themselves.[134] The only option would be for Congress to amend IGRA in a way that allows for general injunctive relief under *Ex parte Young*.[135]

C. Justice Stevens' Dissent

In the first of two lengthy dissents,[136] Justice Stevens attacked the "shocking character of the majority's affront to a coequal branch of government."[137] Stevens endeavored to show that *Chisholm* and *Hans,* the foundational cases for what he derisively calls the "two Eleventh Amendments," were based on interpretations of acts of Congress, not on whether Congress lacked the power to authorize a suit against a state.

To Stevens, the correct reading of the Eleventh Amendment comes from its text and the circumstances surrounding its adoption. The amendment was a reaction to *Chisholm v. Georgia*,[138] in which the Court subjected Georgia to an assumpsit action by a South Carolina creditor brought in federal court.[139] In a famous dissent, Justice Iredell argued that the Judiciary Act of 1789[140] incorporated the common law doctrine of state sovereign immunity, and,

---

129. *Id.* at 1131 ("In putting forward a new theory of state sovereign immunity, the dissent develops its own vision of the political system created by the Framers . . .").

130. *Id.* at 1145 (Souter, J., dissenting) ("[T]he Court today holds for the first time . . . that Congress has no authority to subject a State to the jurisdiction of a federal court at the behest of an individual asserting a federal right.").

131. *Id.* at 1133.

132. *Id.* at 1132.

133. *Id.*

134. *See supra* notes 108-09 and accompanying text.

135. *Seminole Tribe*, 116 S. Ct. at 1133.

136. The dissents in *Seminole Tribe* run 52 pages to the majority's 14 in the commercial reporters. Souter's dissent was the lengthiest opinion of the 1995-96 term.

137. *Id.* at 1134.

138. 2 U.S. (2 Dall.) 363, 419 (1793) (Georgia could not invoke sovereign immunity in an assumpsit case brought by a South Carolina citizen.).

139. *Id.* at 428.

140. Ch. 20, 1 Stat. 73 (1789).

therefore, the federal court had no jurisdiction.[141] Stevens emphasized that Iredell did not argue that the Article III of the Constitution prevented Congress from restricting state sovereign immunity, only that Congress had not done so in the Judiciary Act.[142] The Eleventh Amendment, Stevens argued, was limited to citizen-state diversity cases such as *Chisholm* and did not codify Iredell's expansive view of state sovereign immunity.[143] Thus, Stevens argued, when the Eleventh Amendment was adopted five years after *Chisholm*, it was a partial bar to federal jurisdiction.[144]

Similarly, Stevens argued that the Court in *Hans,* which he sarcastically credits with creating a second Eleventh Amendment, based its opinion on Congressional action, not a constitutional rule of law. The Congressional action in the *Hans* opinion on which Stevens focused was actually inaction: Congress failed to displace the common law presumption of state sovereign immunity. Thus, Stevens concluded that *Hans* established "a presumption against jurisdiction that Congress must overcome, not an inviolable jurisdictional restriction that inheres in the Constitution itself."[145] Stevens reaffirmed the *Union Gas* position that Congress, through the Commerce Clause, may overcome such a presumption[146] and argued the majority's extension of the Eleventh Amendment was incorrect. Stevens' reading is that state sovereign immunity, while affecting federalism, is "subordinate to the plenary power of Commerce."[147]

D. Justice Souter's Dissent

The case's second dissent pursued an exhaustive analysis of state sovereign immunity. Justice Souter's survey of English and American history and precedent runs from the Thirteenth Century reign of Henry III to the present term. He analyzed whether: (1) states enjoyed sovereign immunity from suits brought in their own courts prior to the Constitution; (2) after ratification, that immunity extended to diversity and federal question cases; and (3) Congress could abrogate that immunity.[148]

Souter found no definite answer to the first question,[149] but, by pointing to the ambiguity, he undercut the majority's reliance on *Hans,* which assumed

---

141. *Chisholm*, 2 U.S. at 434-36.

142. *Seminole Tribe*, 116 S. Ct. at 1135.

143. *Id.* at 1136.

144. *Id.* ("Whatever the precise dimensions of the Amendment, its express terms plainly do not apply to all suits brought against unconsenting States.").

145. *Id.* at 1138.

146. *Id.* at 1142 (Stevens argued the Court read the "ancient doctrine of sovereign immunity" into the Eleventh Amendment.).

147. *Id.*

148. *Id.* at 1145.

149. *Id.*

without qualification that states enjoyed sovereign immunity before 1787.[150] For the second, Souter, like Stevens, concluded the Eleventh Amendment, read in the context of the *Chisholm* controversy, granted immunity to states only in citizen-state diversity cases and did not affect federal question jurisdiction.[151] Souter also criticized the *Hans* court for "erroneously" assuming "a State could plead sovereign immunity against a noncitizen suing under federal question jurisdiction, and for that reason held that a State must enjoy the same protection in a suit by one of its citizens."[152] Finally, Souter answered the third question by arguing that the Framers were suspicious of common law doctrines and made them subject to legislative amendment.[153]

According to Souter, in applying this view to the present case, the explicit abrogation of state sovereign immunity to a suit arising from an Indian gaming statute rooted in Article I powers (the Indian Commerce Clause) was trumped by the vague and erroneous notion of sovereign immunity contained in the *Hans* doctrine. "[I]n holding that a non textual common-law rule limits a clear grant of congressional power under Article I, the Court follows a course that has brought it to grief before in our history, and promises to do so again."[154] Souter also argued that the majority had no valid basis for holding that the Seminoles could not obtain prospective injunctive relief under *Ex parte Young*.[155]

III. SEMINOLE TRIBE'S EFFECT ON CERCLA

Arguably, in no other statutory scheme outside of the Indian Gaming Regulatory Act will the fallout of *Seminole Tribe* be more acutely manifest than in CERCLA.[156] CERCLA extends to private parties two substantial mechanisms—private enforcement actions and cost recovery actions—that appear constitutionally inapplicable to unconsenting state defendants. These provisions are the bedrock of CERCLA and their future after *Seminole Tribe* is in doubt because: (1) *Seminole Tribe's* overruling of *Union Gas* gives courts clear guidance on private CERCLA claims brought against states; (2) unlike other environmental regulations,[157] CERCLA is a comprehensive scheme that expressly subjects governmental entities and private parties to liability and is, therefore, not as susceptible to "cooperative federalism;" (3) CERCLA provides a remedial scheme for enforcement of citizen suits, which now are invalid against state agencies (but might be extended to state officials under *Ex parte Young* as interpreted in *Seminole Tribe*); (4) CERCLA specifically provides for cost recovery and contribution actions against states, but with the demise of *Union*

---

150. *Hans*, 134 U.S. at 1.
151. *Seminole Tribe*, 115 S. Ct. at 1145, 1146.
152. *Id.* at 1146.
153. *Id.*
154. *Id.*
155. *Id.*
156. 42 U.S.C. §§ 9601-9675 (1994).
157. *See infra* note 228 and accompanying text.

Gas, the Court's interpretation of the Eleventh Amendment simply invalidates such actions; and (5) the efficacy of CERCLA depends on voluntary settlements, and now that states are in a greater negotiating position with respect to private parties, states have a reduced incentive to settle claims.

A. CERCLA in General

Congress enacted CERCLA in 1980 in response to the public outcry from environmental disasters such as Love Canal in western New York.[158] At the time, existing environmental regulations such as the Resource Conservation and Recovery Act (RCRA)[159] and traditional common law doctrines provided only limited relief. CERCLA's purpose was to provide federal authority to clean up leaking, inactive, or abandoned waste sites and to provide emergency response to spills.[160] The Act makes all potentially responsible parties (PRPs) liable for the clean-up costs on a restitution theory that those responsible for causing hazards should pay.[161] CERCLA's primary goal is to protect and preserve public health and the environment[162] by reaching voluntary settlements.[163] State and local governments continue to play key roles as owners and operators of

---

158. *Hazardous and Toxic Waste Disposal Field Hearings: Joint Hearings before the Subcomms. on Environmental Pollution and Resource Protection of the Comm. on Environment and Public Works*, 96th Cong. 2 (1979) (statement by Sen. Moynihan):

> It was here at the Love Canal that an incident caused by the dumping of hundreds of hazardous chemicals became the first pollution problem recognized to be a national calamity; an unfortunate harbinger of a problem that was at that point just beginning to be understood. It is here that hundreds of families have suffered immense dislocation from their daily lives, in having to move their homes, and in living with growing anxiety over the possible health effects that may be the result of this tragedy. It is here, finally, that we have involvement at all levels of government to address a problem that transgresses all artificial boundries [sic] between local, state, and federal concern.

159. 42 U.S.C. §§ 6901-6992 (1994).

160. *Id.*

161. *See* Lone Pine Steering Committee v. EPA, 777 F.2d 882, 886 (3d Cir. 1985) ("Congress empowered the EPA to take clean up action when necessary," and then collect from responsible parties."). For an excellent overview of CERCLA, see *The Impact of Environmental Law on Real Estate and Other Commercial Transactions*, ALI-ABA Course of Study, Sept. 25, 1997. The author used its fine compilation of cases.

162. *See* Voluntary Purchasing Groups, Inc. v. Reilly, 889 F.2d 1380, 1386 (5th Cir. 1989).

163. *See* United States v. Conservation Chemical Co. 628 F. Supp. 391, 403 (W.D. Mo. 1985), *order modified on other grounds*, 681 F. Supp. 1394 (W.D. Mo. 1988) ("CERCLA contemplates that hazardous waste sites will be cleaned up in the most cost-effective manner; spending precious Superfund monies on a site when there are responsible parties ready and willing to spend private monies to accomplish the same result would hardly be an efficient use of government resources."). *See also* H. REP. No. 96-1016, Part I, at 5 (1980), *reprinted in* 1980 U.S.C.C.A.N. 6119, 6120 (purpose of the bill is to "induce such persons voluntarily to pursue appropriate environmental response actions").

contaminated waste sites, and, as such, they are integral parties in Congress' remedial efforts.[164]

A prime component of CERCLA is the Superfund, a revolving fund created with revenue from a tax on sales of petroleum and other chemicals. Initially, Congress allocated $1.5 billion to the Superfund. The amount was increased to $8.5 billion by the Superfund Amendments and Reauthorization Act of 1986 (SARA).[165] SARA was changed significantly and most prominently by the addition of express contribution cause of action to recover clean-up costs.[166] Another reauthorization vote is likely in the 105th Congress.[167]

B. Structure and Scope of CERCLA and its Relation to States

Federal district courts have "exclusive original jurisdiction over all controversies arising under this chapter, without regard to the citizenship of the parties or the amount in controversy."[168] Parties subject to CERCLA include "the owner and operator of a vessel or facility,"[169] or "any person who at the time of disposal of any hazardous substance owned or operated any facility at which such hazardous substances were disposed of."[170] "Owner or operator" is defined as "any person" who owned or operated a vessel or facility.[171] "Person" means private individuals and entities, the federal government and states, including political subdivisions of states.[172] States are excluded from liability if they are lawfully acting in response to an emergency[173] or if they involuntarily acquire hazardous waste sites.[174] However, such exclusions do not apply to states that cause the environmental damage.[175]

The standard of liability for CERCLA is strict liability patterned after the Clean Water Act.[176] Moreover, courts have ruled that when two or more parties are liable they are joint and severally liable.[177] CERCLA provisions apply

---

164. *See* Brief for Respondent at 8, *Union Gas*, 491 U.S. 1 ("The EPA has estimated that over 16% of all contamination sites on the National Priorities List are currently owned or controlled by states and local governments." Estimates were as of July 1, 1988.).

165. Pub. L. 99-499, 100 Stat. 1613 (1986).

166. Pub. L. 99-499, Title I, § 113, 100 Stat. 1647 (1982) (codified as 2 U.S.C. § 9613 (1994)).

167. CERCLA provision were pending as the 105th Congress recessed for 1997.

168. 42 U.S.C. § 9613(b) (1994).

169. *Id.* § 9607(a)(1).

170. *Id.* § 9607(a)(2).

171. *Id.* § 9601(20)(A).

172. *Id.* § 9601(21).

173. *Id.* § 9607(d)(2).

174. *Id.* § 9601(20)(D).

175. *Id.*

176. *Id* § 9601(32) (specifically cites provisions of the Clean Water Act, 33 U.S.C. §§ 1321-1387 (1994)).

177. *See* United States v. R.W. Meyer, Inc. 889 F.2d 1497, 1507 (6th Cir. 1989) ("CERCLA

retroactively even if the activity in question occurred prior to CERCLA's enactment in 1980.[178]

C. Relevant CERCLA Causes of Action

1. Remediation.—The president and the Environmental Protection Agency (EPA) are given broad authority to initiate and administer large-scale cleanup efforts as part of the National Contingency Plan (NCP).[179] The federal government also may pursue short-term "removal" actions aimed at providing immediate relief at hazardous waste sites. Both remediation and removal actions are limited to sites chosen for the National Priority List, a part of the NCP.[180] The EPA may do the work itself and then recover costs from PRPs, or it may order the PRPs to clean up the site.[181] The EPA may issue orders or seek court-ordered injunctive relief.[182]

2. Cost Recovery and Contribution.—The federal and state governments may pursue cost recovery actions against PRPs.[183] Private parties may do the same, but they may not recover attorney fees.[184]A potentially responsible party can seek contribution from any other PRP liable under CERCLA.[185] This provision is a codification of the common law contribution doctrine that was enforced judicially before Congress amended § 9613 in 1986.[186] "Persons" who resolve their liability to federal or state governments in an administrative or judicially approved plan are not liable for contribution for matters contained in the settlement.[187]

3. Citizen Suits.—Any "citizen" or state may sue the federal government or any other "person" alleged to be in violation of a CERCLA "standard, regulation, condition, requirement, or order."[188] Injunctive relief and civil penalties are available; damages are not.[189] For a person to demonstrate standing to bring a citizen suit, the person must prove an injury in fact.[190]

---

has been interpreted to impose joint and several liability when the environmental harm is indivisible.").

178. *See* United States v. Alcan Corp., CIV No. 95-7570, 1996 WL 489731 (3d Cir. Aug. 22, 1996); *Cf. Olin*, 927 F. Supp. 1502 (S.D. Ala. 1996), *rev'd*, 107 F.3d 1506 (11th Cir. 1997).

179. 42 U.S.C. § 9605(a) (1994).

180. *Id.* § 9604(a).

181. *Id.*

182. *Id.*

183. *Id.* § 9607(a).

184. *See* Key Tronic Corp. v. United States, 511 U.S. 809, 821 (1994) (fees incurred during negotiations with EPA are not "necessary costs of response" and are not recoverable).

185. 42 U.S.C. § 9613(f)(1) (1994).

186. *See Union Gas*, 491 U.S. at 20, 21.

187. 42 U.S.C. § 9613(f)(2) (1994).

188. *Id.* § 9659(a)(1).

189. *See* Regan v. Cherry Corp. 706 F. Supp. 145, 148 (D.R.I. 1989).

190. *See* Conservation Law Foundation of New England Inc. v. Reilly, 950 F.2d 38, 41 (1st

D. Effect of Seminole Tribe on Cost Recovery, Contribution Actions

With the demise of *Union Gas*, states may invoke sovereign immunity in contribution actions, just as Pennsylvania attempted in *Union Gas*. This turns the comprehensive cost-recovery scheme envisioned by Congress—and its value to voluntary settlements—on its head in cases involving states. The ruling violates two major policy objectives of CERCLA: (1) efficient shifting of costs to the parties responsible for pollution and (2) pursuit of cost-efficient settlements, which ultimately will result in faster and more efficient clean-up efforts. By eroding the comprehensive nature of CERCLA, the Court has returned the nation to the pre-1980 morass in which environmental disasters lingered because of inadequate remedies under RCRA, other environmental regulations, and common law principles.[191] In Superfund cases involving states, this change will prolong or eliminate settlements, increasing the cost to taxpayers. Most significantly, however, the change leaves private PRPs with a federal right of action and no court in which to pursue it.

1. Immediate Impact of Seminole Tribe.—Shortly after the ruling in *Seminole Tribe*, two federal district courts dismissed CERCLA claims against state defendants for lack of subject-matter jurisdiction. In *Ninth Avenue Remedial Group v. Allis Chalmers Corp.*,[192] an association of corporations liable under CERCLA brought a contribution action against the Indiana Department of Transportation seeking contribution. The court, following *Seminole Tribe's* rule, recognized the state's Eleventh Amendment immunity and dismissed the suit.[193] The court reasoned that "the fact that CERCLA includes the states as possible liable persons in CERCLA actions filed by private citizens is meaningless in light of" *Seminole Tribe*.[194] It concluded that "any language in CERCLA that makes a state liable to private parties is unenforceable."[195] Plaintiffs in the case failed to persuade the court that the state had waived Eleventh Amendment immunity: (1) through judicial action by virtue of a decision by the Indiana Supreme Court, which held the state had waived immunity in general to tort actions filed in state courts;[196] (2) through statute by incorporating CERCLA definitions in state hazardous substance response provisions; or (3) through the state's conduct in filing CERCLA claims in unrelated lawsuits elsewhere in the state.[197]

In the second case, *Prisco v. New York*,[198] a private owner sought a

---

Cir. 1991) (watchdog group did not have standing for national injunctive relief, but individual members had standing to seek action pertaining to CERCLA sites that actually harmed them).

191. *See supra* note 159 and accompanying text.

192. 962 F. Supp. 131 (N.D. Ind. 1997).

193. *Id.* at 136.

194. *Id.* at 134.

195. *Id.*

196. Campbell v. State, 284 N.E.2d 733, 737 (Ind. 1972).

197. *Id.* at 133-35.

198. 1996 WL 596546, at *14 (S.D.N.Y. Oct. 16, 1996).

declaratory judgment against New York, declaring the state liable under CERCLA for future cleanup costs that will be incurred in a remediation action. The owner alleged that hazardous substances were released on her property while under state control. The district court initially denied the state's motion for summary judgment, holding a material issue of fact remained regarding whether the state exerted control over the property.[199] But after the *Seminole Tribe* ruling, the court granted the state's motion to dismiss the CERCLA claim. From the *Ninth Avenue* and *Prisco* courts' lock-step adherence to *Seminole Tribe*, one can infer that all private CERCLA claims against unconsenting states, including those filed before *Seminole Tribe*, will meet a similar fate. Like results have occurred in other federal cases based on non-Fourteenth Amendment claims.[200]

2. Forum Shifting.—These rulings affirm that CERCLA plaintiffs who oppose states have no cause of action in federal court. CERCLA provides an exclusive federal remedial scheme.[201] Now that private parties cannot bring federal suits against unconsenting states, with the exception of Fourteenth Amendment claims, a logical option would be to pursue some sort of relief in state courts. However, for the following reasons, those options are limited and do not come close to the relief granted by CERCLA before *Seminole Tribe*.

If *Prisco* signals things to come, *Seminole Tribe* sets federal environmental policy back to its pre-CERCLA days. Without a comprehensive remedial scheme authorizing contribution actions against *all* responsible parties, CERCLA is limited severely. This is precisely the danger of which Justice Brennan warned in his majority opinion in *Union Gas*:[202] "If states, which comprise a significant class of owners and operators of hazardous-waste sites . . . need not pay for the costs of cleanup, the overall effect on voluntary cleanups will be substantial." With finite federal resources, voluntary settlement agreements are crucial. Contribution actions are a significant inducement for those settlements. These CERCLA provisions, as Brennan noted, set the act apart from prior non-comprehensive efforts such as RCRA, which "failed in large part because they focused on preventative measures to the exclusion of remedial ones."[203]

On a more practical level, private entities involved in CERCLA actions are harmed significantly by *Seminole Tribe*. For example, "A" is a private party owning a contaminated waste site. The United States (U.S.) initiates a remediation action against A under CERCLA. The U.S. cleans up the site and successfully pursues a cost recovery action against A. Faced with staggering liability, A seeks contribution, under CERCLA, from prior owners and operators of the site. These include U.S., private parties "B," "C," and "D," and A's home state (S). Under CERCLA, any current or prior owner or operator is jointly and severally liable. The U.S. waived sovereign immunity by CERCLA's terms. S has not waived its sovereign immunity and moves for dismissal of the

---

199. 902 F. Supp. 400, 406 (S.D.N.Y. 1995).

200. *See supra* notes 108-09 and accompanying text.

201. 42 U.S.C. § 9613(b) (1994).

202. 491 U.S. at 22.

203. *Id.* at 21.

contribution action against it. Under the ruling in *Seminole Tribe*, A might receive contribution from U.S., B, C, and D, but not from S. Even though CERCLA mandates the parties be strictly liable to each other, the whole system breaks down. This result is especially unfair if S caused significantly more environmental damage than the other responsible parties.[204]

E. Effect of Seminole Tribe on Private Enforcement Through Citizen Suits

Unlike the undressing cost-recovery actions took, CERCLA's citizen-suit provisions survived *Seminole Tribe* relatively unscathed. Although § 9659 expressly grants injunctive relief and civil penalties against any "person" in violation of CERCLA including states,[205] such efforts against states are barred by the Eleventh Amendment as interpreted in *Seminole Tribe*.[206]

What rescues these provisions, however, is the Court's ambiguous reading of the *Ex parte Young* doctrine, which would allow similar actions to be brought against state officials acting in their official capacity. This doctrine holds that state officials are subject to action in federal court for violations of federal law.[207] Although the Court held in *Seminole Tribe* that *Ex parte Young* relief was unavailable to the Seminoles because the federal right in question arose from legislation that included a detailed remedial scheme,[208] the Court left open the possibility that *Ex parte Young* relief was available in regulations with limited injunctive remedies.[209] The Court referenced an environmental regulation, the Clean Water Act, as a limited scheme.[210]

Two federal court rulings following *Seminole Tribe* provide strong evidence of the availability of *Ex parte Young* relief under CERCLA. In *Natural Resources Defense Council v. California Department of Transportation*,[211] an environmental watchdog group sought California's compliance with the Clean Water Act in connection with state management of stormwater runoff from roads and maintenance yards. The state director of transportation also was named in the suit. The district court dismissed the claims against the state but refused to do the same with the claims against the transportation director.[212] The Ninth Circuit, relying on *Seminole Tribe*, held that the environmental watchdog group's claim against the director was valid under *Ex parte Young*.[213] The court reasoned that "Congress implicitly intended to authorize citizens to bring *Ex parte Young*

---

204. For example, if S buried the waste that actually caused the damage.

205. 42 U.S.C. § 9659(a)(1) (1994).

206. *Seminole Tribe*, 116 S. Ct. at 1132.

207. *Ex parte Young*, 209 U.S. at 159-60.

208. *Seminole Tribe*, 116 S. Ct. at 1133.

209. *Id.* at 1133 n.14 (citing provisions in the Clean Water Act 33 U.S.C. 1365(a) (1994) (relief is available against "any person").

210. *Id.*

211. 96 F.3d 420 (9th Cir. 1996).

212. *Id.* at 423.

213. *Id.*

suits against state officials with the responsibility to comply with clean water standards and permits."[214]

Significantly, CERCLA's citizen suit provisions[215] track those of the Clean Water Act.[216] It is reasonable to infer that environmental groups seeking CERCLA compliance from states will be able to pursue *Ex parte Young* actions by naming state officials as the Natural Resources Defense Council did in the Ninth Circuit. The district court in *Prisco* makes the same point in dicta.[217] The court allowed the property owner to maintain an *Ex parte Young* action against New York for claims seeking injunctive relief under RCRA,[218] which has a citizen suit provision similar to CERCLA's.[219] The *Prisco* court, in a footnote, distinguished the private owner's CERCLA claim with respect to *Ex parte Young* because the CERCLA claim sought damages.[220]

IV. REMEDIES AND ACCOMMODATIONS

The harm done by *Seminole Tribe* to CERCLA cannot be remedied easily. The goal of any strategy to mitigate the harshness of the decision should be to restore the balance between the national goals of environmental health and safety and the rights of states. State sovereignty should be respected, but only so far as it promotes state accountability. Some suggested remedies and accommodations follow:

A. Total Remedies: Constitutional Amendment or Reversal Seminole Tribe

1. Constitutional Amendment.—As farfetched as this solution appears, it is not without precedent. A recasting of the Eleventh Amendment could end any debate about the scope of state sovereign immunity to federal suits. Provisions could be added that expressly subject states to federal jurisdiction in federal question cases. Contrarily, an amendment could codify the *Hans* doctrine, which holds that states enjoy sovereign immunity from all federal suits brought by private parties.[221] The first suggested redrafting would effectively further the goals of federalism, making states accountable when national interests are at stake, as in CERCLA cases.

When faced with puzzling federalism questions, the nation's leaders more

---

214. *Id.* at 424. The court is referring to 33 U.S.C. § 1365(a)(1) (1994) (citizen suits may be brought against "any person" who violates relevant provisions).

215. 42. U.S.C. § 9659(a)(1) (1994) (CERCLA's citizen suit provision also authorizes "any person" to bring such a suit).

216. 33 U.S.C. § 1365 (1994).

217. *Prisco*, 1996 WL 596546, at *16.

218. *Id.*

219. 42 U.S.C. § 6972(a)(1)(B) (1994)(Citizen suits are permitted against any "governmental instrumentality or agency" as allowed under the Eleventh Amendment.).

220. *See Prisco*, 1996 WL 596546 at *15 n.24.

221. Although it can be readily assumed that such provisions would be drafted as to not counter the Fourteenth Amendment.

than once have amended the Constitution to correct the federalism balance.[222] One Supreme Court case, *Chisholm*, led directly to the adoption of the Eleventh Amendment. However, nothing close to the national debt crisis that colored *Chisholm* exists today. If the Court had struck down abrogation under the Fourteenth Amendment, and civil rights enforcement had been brought into question, the public reaction probably would have been greater. Absent such a development, the Eleventh Amendment appears safe as a second-shelf provision. But recent efforts to amend the Constitution in the name of flag preservation, abortion restrictions, and budget balancing demonstrate that a federalism-inspired change should not be ruled out.

2. Reversing Seminole Tribe.—Reversing the ruling would not be much easier. Each faction in the federalism split could lose a justice—Chief Justice Rehnquist in the majority and Justice Stevens in the dissent—to retirement in the next few years, so the effect of a lineup change could be nil. Assuming no changes, the swing votes on the court, Justices Kennedy and O'Connor, appear firmly in the *Seminole Tribe* majority's camp concerning the Eleventh Amendment. Unlike *Lopez,* which established a test of degrees that, in another setting, could send Justice Kennedy to the liberal bloc,[223] *Seminole Tribe* did not leave much of a gray area. Justice Kennedy sustained the judgment in *Union Gas*,[224] but not its reasoning, and he agreed with the basic holding of *Seminole Tribe*—that Congress cannot rely on Commerce Clause authority to abrogate state sovereign immunity to private federal suits. Therefore, a favorable lineup change and continued solidarity are the *Seminole Tribe* dissent's best chance for a quick turnabout.

B. Middle Ground Test Between Seminole Tribe and Union Gas

When the Supreme Court hears its next Article I abrogation case, it should consider a compromise position first put forth in an amicus curiae brief in *Union Gas*.[225] Rather than decide abrogation as an all-or-nothing proposition, as the Court did in *Union Gas* and *Seminole Tribe*, the Court could adopt a three-prong test that would narrow abrogation to important national interests. Under the test,

---

222. Several constitutional amendments owe their existence to federalism controversies: Tenth (residual powers vested in states), Eleventh (states immune to private suits in federal court), Thirteenth (slavery abolished), Fourteenth (extensive personal rights including due process of law), Fifteen (voting rights protected), Sixteenth (Congress may lay income tax), Seventeenth (direct election of senators). U.S. CONST. amend. X, XI, XIII-XVII.

223. *See* Singer, *supra* note 67.

224. *Union Gas*, 491 U.S. at 23 (Congress may abrogate state sovereign immunity under the Commerce Clause. This was overruled by *Seminole Tribe*, 116 S. Ct. at 1128.)

225. *Union Gas*, 491 U.S. 1. Amicus Curiae Brief of Pacific Legal Foundation in Support of Respondent. *See also*, Flournoy v. California, 230 Cal. App. 2d 520, 537 (1964) (The court applied the three-part policy test to find a California tort claims statute, 1963 Cal. Stat. ch. 1681 § 45, applied retroactively in a suit brought by an injured motorist against the state for alleged negligent bridge construction.).

the Court would consider (1) the nature and strength of the policy interest served by the statute; (2) the extent to which the statute abrogates the Eleventh Amendment; and (3) the nature of the right that the statute alters.[226] In applying these factors to a CERCLA case such as *Union Gas*, the Court would first find the nature and strength of the policy interest to be the protection of public health from known dangers. Second, CERCLA only abrogates immunity to contribution actions in which the state is a responsible party. CERCLA does not supplant all of a state's sovereign immunity. Finally, the nature of the right the statute alters is limited. Parties have the right to pursue contribution and cost recovery actions in federal court. This is not a significant blow to states. Private parties will succeed only to the extent they can prove states are liable.

C. Spending Clause Solutions

1. CERCLA Amendments.—A third option includes amending CERCLA to take advantage of the "cooperative federalism" dynamics the Court endorsed in *New York v. United States.*[227] Unlike its role in other environmental regulations, the federal government retains almost all of the authority in CERCLA.[228] Its ambitious twin goals of cleaning up national hazardous waste disasters and making polluters pay[229] leave states in a secondary role.[230] Now that *Seminole Tribe* has strengthened the states' hand, Congress could offer incentives to states conditioned on the states' waiver of sovereign immunity to private CERCLA suits. Such a strategy has the dual benefits of respecting states' autonomy and restoring the comprehensive remedial scheme that existed before *Seminole Tribe*.

Congress has the authority, grounded in the Spending Clause,[231] to insist that states comply with federal regulations as a condition of receiving federal

---

226. For this test to be constitutional, the Court first would have to overturn *Hans*, 134 U.S. 1, which barred all private federal suits against unconsenting states. The Court would then have to interpret the Eleventh Amendment as not barring federal-question jurisdiction in such suits.

227. 505 U.S. at 167 (Congress may, under its power to regulate private activity under the Commerce Clause, offer states the choice of regulating the activity according to federal standards or having state law pre-empted by federal regulation.).

228. Under other regulations such as the Resource Conservation Recovery Act, 42 U.S.C. §§ 9601-9675 (1994), and the Low-Level Radioactive Waste Policy Act, 42 U.S.C. §§ 2021b-2021j, (1994), states play a more active role in implementing federal policies than states do under CERCLA. *See New York*, 505 U.S. 144, 151 (Radioactive waste regulation depends on interstate compacts). For an excellent discussion of "cooperative federalism" in the environmental context, see two articles from a 1995 symposium at the University of Maryland. Robert V. Percival, *Environmental Federalism: Historical Roots and Contemporary Models*, 54 MD. L. REV. 1141 (1995); Adam Babich, *Our Federalism, Our Hazardous Waste, and Our Good Fortune*, 54 MD. L. REV. 1516 (1995).

229. *See supra* Part III.A.

230. For example, states may initiate cleanups under CERCLA and then recover their costs from potentially responsible parties. 42 U.S.C. § 9607 (1994).

231. U.S. CONST. art I, § 8, cl. 1. *See supra* note 74 and accompanying text.

funding.[232] The federalism concern in such an undertaking is whether the condition on states is a constitutionally valid incentive or an unconstitutional coercion.[233] Furthermore, the condition has to satisfy factors articulated by the Court in *South Dakota v. Dole* to be considered valid under the Spending Clause. Those factors include (1) the exercise of Spending Clause power must be in pursuit of "the general welfare;" (2) the condition must be unambiguous; (3) the conditions must relate to a national interest; and (4) it must not be proscribed by other constitutional prohibitions.[234]

Congress should have no trouble complying with these requirements in a CERCLA amendment. To protect the condition as an incentive, not coercion, Congress must comport with the Court's findings in *New York v. United States*[235] by offering states a meaningful choice. Congress should be able to do that, as discussed below. The *Dole* factors should not pose a problem. First, the general welfare connection is clear—national public health and safety. Second, express terms easily could make the conditions unambiguous. Third, by eradicating dangerous and costly waste sites, the national interest is promoted. Finally, no other constitutional provisions bar the conditioning of such funds as long as Commerce Clause and Spending Clause interpretations are met.

Specifically, Congress should amend § 9607(a)(4), which allows the federal and state governments[236] and private individuals[237] to recover response costs. The grant of the cause of action to states in subparagraph A should be redrafted to say "any State that has waived its sovereign immunity for claims filed under this chapter" may recover response costs. Then, subparagraph B should be amended to close the loophole of "any person," in which a state could fall based on the chapter definition of "person."[238] This latter amendment could be accomplished by making subparagraph B subject to subparagraph A.

States that want to preserve sovereign immunity would do so at great cost. The cost-recovery actions are the most effective mechanism under CERCLA for states to shift costs to polluters. However, it would be fundamentally unfair to offer a state, a potential wrongdoer under CERCLA, the benefits of cost recovery in this comprehensive scheme without imposing an equitable cost. The price the state must pay is subjecting itself to suit brought by private parties. Under §

---

232. *See* South Dakota v. Dole, 483 U.S. 203, 208-210 (1987) (Congress had authority under the Spending Clause to condition receipt of federal highway funding on state's adoption of a minimum age for the purchase of alcoholic beverages).

233. *Id.* at 211.

234. *Id.* at 207-08.

235. 505 U.S. 144, 185 (1992).

236. 42 U.S.C. § 9607(a)(4)(A)(1994) (Persons liable under CERCLA are liable for "all costs of removal or remedial action incurred by the United States Government or a State or an Indian tribe not inconsistent with the national contingency plan.").

237. *Id.* § 9607(a)(4)(B) (Similar to subparagraph A, parties are liable for "any other necessary costs of response incurred by any other person consistent with the national contingency plan.").

238. *Id.* § 9601(21) (Person includes a "State" or "political subdivision of a State.").

9607, a state may clean up a site, pursuant to the National Contingency Plan, and then impose joint and several liability on responsible parties. While the federal government retains oversight, states may play a leading role in cleaning up Superfund sites in their backyards under the auspices of CERCLA. States should want to assume this role out of public policy concerns and self-interest. The public policies forwarded include the furtherance of public health and safety, the enforcement of liability of those who cause damage, and the return of land to productive use. Out of self-interest, states would seek quick, cost-efficient cleanups that maximize the use of state resources. States that do not make use of cost-recovery actions must wait until EPA dictates how the site will be handled, creating the possibility of expensive delays. Also, states that clean up sites and remove them from the Superfund list capture more federal resources than states that do not.[239]

Nonetheless, in keeping with "cooperative federalism" goals, states would have the option of not accepting the right to cost-recovery under § 9607. A state could preserve its sovereign immunity to private suits brought under CERCLA, as *Seminole Tribe* says it may, and recover its costs under state regulations or CERCLA contribution provisions. Both options compare unfavorably to § 9607. State remedies are more limited in scope and may not be pursued if they conflict with CERCLA.[240] If a state files a contribution action under CERCLA[241] in federal court, it waives sovereign immunity to private counterclaimants.[242]

A key part of the cost-recovery-condition plan would be to retain a state's right to contribution under § 9613, so that the constitutionally mandated choice exists.[243] It would be unfair, and likely unconstitutional under the *New York v. United States* framework, to subject a state to severe liability under CERCLA without giving it a means to prove that other PRPs share liability. However, contribution is a more limited remedy because it contemplates liability as between joint tortfeasors. Nonliable parties can bring cost-recovery actions, and defendants can be held jointly and severally liable. Another practical consideration is that the statute of limitations is six years for cost-recovery actions and three years for contribution.[244]

In sum, amending CERCLA to offer states the continued availability of cost-recovery actions in return for their waiver of sovereign immunity to private suits

---

239. For example, under 42 U.S.C. § 9622 (1994), the federal government has wide discretion to provide "mixed funding" financing to facilitate cleanups. States taking advantage of this provision would not have to commit as much of their own resources.

240. *See supra* Part IV.C.1.

241. 42 U.S.C. § 9613 (1994).

242. *See* United States v. Mottolo, 605 F. Supp. 898, 910 (D.N.H. 1984) (A state waives its Eleventh Amendment and sovereign immunities to compulsory recoupment counterclaims by filing a complaint in federal court.).

243. A CERCLA provision, 42 U.S.C. § 9613(f)(1) (1994), preserves this right. ("Nothing in this subsection shall diminish the right of any person to bring an action for contribution in the absence of a civil action under [sections 9606 or 9607].").

244. 42 U.S.C. § 9613 (1994).

may not give states much of a choice, but it is a choice nonetheless. It would restore the balance between state and federal governments needed to address a serious national problem and preserve private parties' rights.

2. Non-CERCLA Inducements.—Congress also could induce states' full participation in CERCLA by requiring a sovereign immunity waiver as a condition to receiving non-CERCLA environmental grant money.[245] The key issue would be whether, under the *Dole* factors, the conditioned spending relates to a national interest.[246] A strong argument could be made that these environmental grants are aimed at promoting and protecting public safety and health by effectively managing waste and toxic substances. CERCLA's goals include the cleanup of mismanaged waste and toxic substances. Thus, a strong relation exists. In addition, states could not claim coercion if they voluntarily decline the grant money. The public policy arguments are the same as those for CERCLA amendments discussed above. States that want to address environmental problems will gain assistance from the federal government and capture dwindling grant money from states who maintain sovereign immunity with respect to CERCLA. Non-participating states that are polluters should pay for their conduct and should be deprived of federal grants. Those grants would go to other states willing to be held fully accountable for their acts. Ultimately, states rejecting the condition will spend more of their own resources, delay the return of waste sites to productive use, and continue to impinge on private parties' rights.

D. State Court Actions

Now that private parties do not have an exclusive federal remedy against unconsenting states under CERCLA, many parties may have to turn to the state courts to seek contribution from state-government defendants. In a sense, this returns prospective plaintiffs to pre-CERCLA days when polluters were subject to common doctrines of nuisance and contribution. However, the comprehensive dynamics of CERCLA have spawned litigation in a dramatically different context. Thus, attorneys who represent private clients in CERCLA contribution actions will tread new ground pursuing related claims in state courts. Unique problems may arise related to partial state sovereign immunity, the lack of expertise of state judges, and state political pressures.

1. State Tort Claims Acts.—The initial consideration for an attorney in this setting is finding a court of competent jurisdiction by drawing on the text of CERCLA, the ruling in *Seminole Tribe*, and assessing the scope of state

---

245. Examples include grants that help states construct wastewater treatment plants, 40 C.F.R. Solid § 35.300 (1995); maintain public drinking water systems, 40 C.F.R. § 35.400 (1995); implement hazardous waste management programs, 40 C.F.R. § 35.500 (1995); and develop non-point groundwater pollution control plans, 40 C.F.R. § 35.750 (1995).

246. Assuming other factors would be met: the activity promotes the public welfare, the condition is clear, and no other constitutional provision bars the condition. *See Dole*, 483 U.S. at 207-08.

sovereign immunity in that particular state. In pre-*Seminole Tribe* CERCLA cases in which parties sought partial relief under state tort claims acts, courts ruled that CERCLA preempted such actions.[247] However, the preemption of state tort claims acts only extends to cases in which the state statute conflicts with federal law.[248] Now that the provisions of CERCLA relating to private enforcement of rights against states appear unconstitutional, courts should allow plaintiffs to recover in CERCLA-related actions under state tort claims acts. The theory of recovery would be that the state statutes no longer conflict with certain CERCLA provisions rendered unconstitutional by the Court's interpretation in *Seminole Tribe*. Non-conflicting state law claims are valid in CERCLA-related litigation.[249]

Once the decision is made to pursue a state court claim, the attorney should assess that particular state's sovereign immunity doctrines. All states have curtailed sovereign immunity by judicial action or by statute.[250] The idiosyncrasies of each statute or state rule will affect litigation strategy. An early CERCLA case, *Artesian Water Co. v. New Castle County*,[251] illustrates how state tort claims acts might bar state suits. In *Artesian*, a private water company filed a contribution claim against a county. The county argued that the contribution claim failed because the county was protected by the Delaware Tort Claims Act,[252] which expressly protected governmental units from damage claims arising out of the release of pollutants into bodies of water.[253] Ruling eleven years before *Seminole Tribe*, the court in *Artesian* dispatched the county's claim on grounds that CERCLA, pursuant to the Supremacy Clause,[254] preempted the state

---

247. *See* Artesian Water Co. v. New Castle County, 605 F. Supp. 1348, 1354 (D. Del. 1985); United States v. Seymour Recycling Corp., 686 F. Supp. 696, 700 (S.D. Ind. 1988).

248. *Artesian*, 605 F. Supp. at 1354.

249. 42 U.S.C. § 9614(a) (1994) ("Nothing in this chapter shall be construed or interpreted as preempting any State from imposing any additional liability or requirements with respect to the release of hazardous substances within such state."). *See* Attorney General v. Thomas Solvent Co., 380 N.W.2d 53, 59-60 (Mich. Ct. App. 1985) ("It is clear that CERCLA was intended only to supplement hazardous waste programs and not to preempt state programs.").

250. For a comprehensive list of the status of sovereign immunity in all states and the District of Columbia, see RESTATEMENT (SECOND) OF TORTS § 895B app. at 256-270 (1977), app. at 252-56 (1979 & Supp. 1996); *see also* W. PAGE KEETON ET AL, PROSSER AND KEETON ON THE LAW OF TORTS § 131 at 1043-56 (5th ed. 1985 & Supp. 1988).

251. 605 F. Supp. 1348 (D. Del. 1985).

252. 10 DEL. CODE ANN. §§ 4001-4013 (Supp. 1996).

253. *Id.* § 4011(b)(5) ("The discharge, dispersal, release or escape of smoke, vapors, soot, fumes, acids, alkalines, toxic chemicals, liquids or gases, waste materials or other irritants, contaminants or pollutants into or upon land, the atmosphere or any watercourse or body of water, except as provided in subdivision (3) of § 4012 of this title.").

254. U.S. CONST. art VI, cl. 2, which provides:

> This Constitution, and the Laws of the United States which shall be made in Pursuance thereof; and all Treaties made, or which shall be made, under the Authority of the United States, shall be the supreme Law of the Land; and the Judges in every State shall

tort claims act.[255] The court reasoned that the tort claims act conflicted with Congress' intent to hold governments liable under CERCLA.[256]

Today, the validity of the *Artesian* ruling is suspect. The relevant Delaware statutes have not changed, but the *Seminole Tribe* ruling has trumped the CERCLA provisions that the *Artesian* court said preempted the Delaware tort claims act. Plaintiffs such as the water company in *Artesian* can no longer sue unconsenting government defendants in federal court. If the case were brought in Delaware courts, the pollution exemption in the state tort claims act would likely bar the suit. The result is that a party might have a valid contribution claim against a tortfeasor and no forum in which to seek justice. This could possibly give rise to a constitutional claim, with the plaintiff claiming the state deprived it of property rights without due process of law.[257] Delaware's near total retention of sovereign immunity is probably the worst hurdle a plaintiff would face. At the very least, this illustrates the trouble plaintiffs could encounter when taking CERCLA-related claims to state courts.

California—on the opposite end of the spectrum—has a more expansive tort claims act.[258] However, California does not grant the state immunity in important areas affecting CERCLA claims. In *United States v. Montrose Chemical Corp. of California*,[259] a federal district court interpreting California's tort claim act held that the state was not immune to a counterclaim for damages brought by a defendant chemical company.[260] The court found that the chemical company stated several statutory causes of action that fall outside the immunity granted by the tort claims act.[261] The court recognized the chemical company's right to recoupment in state court subject to proving the state's liability.[262] To apply such a ruling today, it is clear that California makes itself amenable to state suits brought by private parties in the context of CERCLA. However, claimants likely would have to prove higher liability standards, such as negligence or nuisance, than the strict liability provided by CERCLA.

Indiana could be considered a middle ground between Delaware's express assertion of immunity to some environmental suits and California's expansive tort claims doctrine. The right to sue the state of Indiana is guaranteed in the

---

be bound thereby, any Thing in the Constitution of Laws of any State to the Contrary notwithstanding.

255. *Artesian*, 605 F. Supp. at 1355.

256. *Id.* at 1354.

257. U.S. CONST. art. V and amend. XIV.

258. CAL. GOV'T CODE §§ 810-997.6 (West 1995 & Supp. 1996).

259. 788 F. Supp. 1485 (C.D. Calif. 1992).

260. *Id.* at 1494.

261. CAL. GOV'T CODE § 815 (West 1995) provides that "[e]xcept as otherwise provided by statute: (a) A public entity is not liable for an injury, whether such injury arises out of an act or omission of the public entity." The *Montrose* court then found that the chemical company stated statutory causes of action under negligence, nuisance, dangerous condition of public property, and failure to discharge mandatory duty. *Montrose*, 788 F. Supp. at 1495.)

262. *Montrose*, 788 F. Supp. at 1494.

state constitution.[263] State and federal courts have ruled that parties can sue the state in its courts subject to statutory limitations and common law.[264] Any CERCLA-related claim would have to get past the several immunities Indiana retains by statute.[265] In a negligence and nuisance suit filed against a town for its alleged faulty construction and maintenance of a sewage system,[266] a state court found that property owners stated a cause of action and rejected the town's claim of immunity based on its discretionary powers.[267] Indiana has not entertained a CERCLA-related case such as *Montrose*.[268] But similar to California, Indiana imposes fault-based liability on governmental entities subjected to suit.[269] Thus, plaintiffs lose the strict liability of CERCLA, but they can seek damages against the state.

2. Practical, Economic, and Political Factors.—A second consideration in pursuing state forums is the expertise of the judges. Federal district judges have tried CERCLA cases exclusively. They are familiar with the nuances and complexities of the cases. Assuming that many state trial judges have handled similarly complicated environmental litigation, it is also highly probable that the range of expertise in state courts is considerably wide. Attorneys would have to compensate by providing more detailed pleadings, take a more proactive role in pre-trial proceedings, and even educate judges on CERCLA.

Thirdly, haling states into their own courts also means that attorneys will have to consider the effects of state politics. Superfund sites are large, expensive, and very public problems. State trial judges will decide their state's liability, which will be paid for by state and local tax money. These judges, some of whom are elected or retained by state and local taxpayers, face political and economic realities that federal judges, with life appointments, do not face.

If states would waive their sovereign immunity out of public policy concerns,

---

263. IND. CONST. art. IV, § 24. ("Provision may be made, by general law, for bringing suit against the State; but no special law authorizing such suit to be brought, or making compensation to any person claiming damages against the State, shall ever be passed.").

264. *See* Burr v. Duckworth, 547 F. Supp. 192, 195 (N.D. Ind. 1982), *aff'd,* 746 F.2d 1482 (7th Cir. 1984) (Indiana no longer adheres to strict sovereign immunity and is not immune from damages resulting from the exercise of its proprietary or governmental functions); Maroon v. Dept. of Mental Health, 411 N.E.2d 404, 415 (Ind. Ct. App. 1980) (the state, as sovereign, may provide conditions and limits for bringing suits against itself); Perkins v. State, 251 N.E.2d 30, 32 (Ind. 1969) (question of state or sovereign immunity rests upon common law).

265. IND. CODE § 34-4-16.5-3 (1993 & Supp. 1996). For example, "A governmental entity . . . is not liable if a loss results from: . . . (6) the performance of a discretionary function . . . ."

266. Hodge v. Town of Kingman, 519 N.E.2d 1266 (Ind. Ct. App. 1988).

267. *Id.* at 1270 (rejecting town's claims under IND CODE § 34-4-16.5-3(6) because actions of town officials were ministerial in nature, not discretionary).

268. *See* United States v. Seymour Recycling Corp., 686 F. Supp. 696, 700 (S.D. Ind. 1988) (In a pre-*Seminole Tribe* ruling, a federal court ruled CERCLA preempted the Indiana Tort Claims Act, and the court therefore did not reach issues of state liability.).

269. *See* Rodman v. City of Wabash, 497 N.E.2d 234, 240 (Ind. Ct. App. 1986) (duty of reasonable care met by municipality to homeowners complaining of sewage backups).

the *Seminole Tribe* holding would be benign. However, many might act as Pennsylvania did in *Union Gas*,[270] Indiana did in *Ninth Avenue*,[271] and New York did in *Prisco*[272] and seek dismissal of claims against them. A third possibility is that a state would yield to suit in its own courts under certain conditions. Attorneys faced with the third scenario should weigh carefully the legal, political, and economic ramifications, as discussed above, in planning a CERCLA-related claim in state courts.

CONCLUSION

The Supreme Court's decision in *Seminole Tribe of Florida v. Florida* upset the balance of federalism, thwarting Congress' efforts to provide an effective remedy to a national crisis. As a result, private parties have seen their rights eroded. The national goal of efficient, voluntary cleanup efforts was dealt a serious blow. No longer may private parties seek contribution claims against unconsenting state defendants. States that caused environmental damage will now be able to shift the cost of their misdeeds to private parties. While the best solution is a repudiation of *Seminole Tribe*, lawmakers, attorneys, and judges should make efforts at the federal and state levels—through amendments to CERCLA, conditioned federal spending, and state judicial actions—to restore the federalism balance.

---

270. *See Union Gas*, 491 U.S. 1.

271. *See Ninth Avenue*, 962 F. Supp. 131.

272. *See Prisco*, 1996 WL 596546.

WILL NASCAR HAVE TO PUT ON THE BRAKES?: THE CONSTITUTIONALITY OF THE FDA'S BAN ON BRAND-NAME TOBACCO SPONSORSHIP IN MOTOR SPORTS

SCOTT D. MATTHEWS*

INTRODUCTION

They come in their motor homes and pick-up trucks, the working class and the wealthy alike. Some come with coolers full of beer. Some come in tee shirts sporting their hero. Some come with no shirts at all. But they all come for the same reason—to see cars blaze around a race track at nearly two-hundred miles per hour; to see their favorite driver bump fenders and "trade paint" with forty other cars on the track; and to see death-defying accidents. When they walk through the turnstiles, the one thing that they are sure to see are advertisements. Welcome to the world of NASCAR auto racing.

Beer, batteries, brakes. Paint, pain reliever, and Pepsi. If you can spell it, they will sell it. On race cars, billboards, flags, tee shirts and caps, advertisements for brand-name products appear everywhere. Above all of this massive marketing barrage, one product stands taller than the others. That product is tobacco, but its days as NASCAR's preeminent sponsor may be short-lived.

On August 23, 1996, President Clinton proclaimed: "With this historic action that we are taking today, Joe Camel and the Marlboro Man will be out of our children's reach forever."[1] The historic act to which the President was referring was the announcement that nicotine is an addictive drug and, as a result, the Federal Food and Drug Administration (FDA) has jurisdiction to regulate tobacco.[2] As a result, the FDA promulgated regulations which govern the access to and promotion of cigarettes and smokeless tobacco to children and adolescents.[3]

The regulations faced their first challenge in a North Carolina District in the

---

* J.D. Candidate, 1998, Indiana University School of Law—Indianapolis; B.A., 1993, Indiana University, Bloomington, Indiana.

1. *Clinton Unveils Tough Tobacco Rules*, INDIANAPOLIS STAR, Aug. 24, 1996, at A1.

2. *See infra* Part II.B. This Note does not address the jurisdictional issue in any depth. *But see* Ann Mileur Boeckman, *An Exercise in Administrative Creativity: The FDA's Assertion of Jurisdiction Over Tobacco*, 45 CATH. U. L. REV. 991 (1996) and Michael Whatley, *The FDA v. Joe Camel: An Analysis of the FDA's Attempt to Regulate Tobacco and Tobacco Products Under the Federal Food, Drug, and Cosmetic Act*, 22 J. LEGIS. 121 (1996) (both addressing the jurisdictional issue prior to the *Coyne Beahm* decision).

3. Regulations Restricting the Sale and Distribution of Cigarettes and Smokeless Tobacco to Protect Children and Adolescents, 61 Fed. Reg. 44,396 (1996) (codified at 21 C.F.R. pts. 801, 803, 804, 807, 820, and 897) [hereinafter FDA Regulations]. The rule regulates only the sale and distribution of cigarettes and smokeless tobacco to adolescents. It does not apply to the sale and distribution of cigars or pipe tobacco.

fall of 1996.[4] The tobacco industry filed suit claiming that the FDA lacked authority to regulate tobacco products. On April 25, 1997, United States District Judge William L. Osteen, Sr. ruled that the FDA had the authority to regulate cigarettes and tobacco products but that it had exceeded its authority by promulgating restrictions and regulations on the advertisement and promotion of tobacco products.[5] Judge Osteen declined to decide whether the FDA's regulations violated the First Amendment "[i]n light of the court's finding that [the] FDA lacks authority . . . to restrict the promotion and advertising of tobacco products."[6] Both sides have appealed the ruling to the Fourth Circuit; however, as of the date of this Note's publication, no opinion has been issued. It is likely that the Fourth Circuit's decision will be ultimately reviewed by the United States Supreme Court.

On June 20, 1997, with the appeal to the Fourth Circuit still pending, the FDA and the tobacco industry reached a proposed settlement agreement that would resolve this issue.[7] The proposed settlement encompasses the FDA rule promulgated on August 28, 1996 and substantially extends it in many aspects.[8] The proposed settlement must still be approved by President Clinton and Congress. If the proposed settlement is approved, the judicial appeal to the Fourth Circuit would be dismissed by the tobacco industry and all advertising restrictions contained in both the FDA rule and the proposed settlement would

---

4. Coyne Beahm, Inc. v. FDA, 966 F. Supp. 1374 (M.D.N.C. 1997).

5. *Id.* at 1400.

6. *Id.* at 1400 n.33.

7. *Tobacco Settlement Reached by the State Attorneys General and the Tobacco Industry*, June 20, 1997 (available on Westlaw at 121 DER T-14) [hereinafter *Tobacco Settlement*].

8. *See infra* Part II.B (discussing the provisions of the advertising restrictions in the FDA Rule). The proposed settlement is much broader than the original rule. Under the agreement, the tobacco industry would pay approximately $368.5 billion over 25 years, including $60 billion in lieu of punitive damages for past conduct. The rights of individuals to sue for compensatory damages would not be abridged and the tobacco industry would be liable for punitive damages for any future conduct. All lawsuits filed against the industry by state attorney generals would be settled in return for a substantial payment that would reimburse states for the tobacco-related costs they have incurred. Furthermore, all lawsuits against the FDA would be dismissed.

In addition to the proposed regulations promulgated by the FDA, the settlement incorporates terms not part of the original FDA rule. With respect to the advertising restrictions, the additional terms include: 1) the elimination of all billboards and outdoor signs; 2) the elimination of all human images and cartoon characters that are used to advertise tobacco products; 3) additional restrictions on point-of-purchase advertisements, including restrictions on their placement in retail stores; 4) the elimination of advertising on the Internet; 5) the ban of direct and indirect payments for tobacco product placement in movies, television programs, and video games; 6) prohibiting direct and indirect payments to "glamorize" tobacco use in the media; and 7) relevant to the subject matter of this Note, protection from a First Amendment challenge. All advertising restrictions contained in the August 28, 1996 FDA rule and the settlement agreement would be placed in consent decrees that would insulate the FDA from constitutional challenges by the tobacco industry and third parties not part of the settlement agreement. *Tobacco Settlement*, *supra* note 7.

be insulated from First Amendment challenges by any party inside or outside of the tobacco industry.

As of publication, neither the President nor Congress has approved the proposed settlement, and it appears that neither ever will. Therefore, the constitutionality of the FDA rule will be left for the courts to decide. Even though Judge Osteen deferred comment on the First Amendment challenge, it is likely that the FDA's rule will be subjected to judicial scrutiny, either by the Fourth Circuit or the Supreme Court, under a First Amendment challenge if the settlement agreement is not accepted.

In a year's time, the FDA has changed its position from claiming that it had no authority to regulate tobacco products, to promulgating sweeping regulations, to negotiating a proposed settlement with the tobacco industry. Why have these recent developments unraveled at breath-taking speed? The reason is simple—smoking has become a major health concern in the United States.[9] Although the sale of tobacco products to persons under the age of eighteen is illegal in every state,[10] most tobacco users begin using tobacco before they reach eighteen.[11] Approximately fifty million Americans smoke cigarettes, while another six million use smokeless tobacco.[12] Four million adolescents smoke cigarettes or use smokeless tobacco,[13] while three thousand persons under the age of eighteen start using tobacco each day.[14] Ninety percent of all new smokers are under the age of eighteen.[15] According to FDA reports, youth smoking is on the rise.[16] Each year, experts estimate that children and adolescents smoke between 516 million and 947 million packages of cigarettes, and use another 26

---

9. Each year, more than 400,000 people in the United States die from tobacco related illnesses. FDA Regulations, *supra* note 3, at 44,398. Tobacco kills more people than AIDS, car accidents, homicides, alcohol, illegal drugs, suicides, and fires combined. *Id.*

10. *Id.* at 44,397.

11. *Id.* at 44,398. Eighty-eight percent of smokers smoked their first cigarette before they were 18 and 71% were daily smokers before reaching 18. Charles J. Harder, *Is it Curtains for Joe Camel? A Critical Analysis of the 1995 FDA Proposed Rule to Restrict Tobacco Advertising, Promotion and Sales to Protect Children and Adolescents*, 16 LOY. L.A. ENT. L.J. 399, 400 (1995) (citing Dep't of Health and Human Servs., Preventing Tobacco Use Among Young People: A Report of the Surgeon General 58 (1994)).

12. FDA Regulations, *supra* note 3, at 44,398.

13. *Id.*

14. Harder, *supra* note 11, at 404 (citing John P. Pierce et al., *Trends in Cigarette Smoking in the United States, Projections to the Year 2000*, 261 JAMA 61, 64 (1989)).

15. Richard Lacayo et al., *Put Out the Butt, Junior*, TIME, Sept. 2, 1996, at 51.

16. Between 1991 and 1994, the number of eighth graders who smoked rose 30%, from 14.3% to 18.6%; among twelfth grade students, 31.2% of the students smoked in 1994, as compared to 28.3% in 1991; smoking increased among college freshmen from 9% to 12.5% during the same time period. Regulations Restricting the Sale and Distribution of Cigarettes and Smokeless Tobacco Products to Protect Children and Adolescents, 60 Fed. Reg. 41,314, 41,315 (proposed Aug. 11, 1995) [hereinafter Proposed Regulations].

million containers of smokeless tobacco.[17]

Under the FDA's final rule, the regulations prohibit, *inter alia*, the "sponsorship of sporting and other events, teams, and entries in a brand name tobacco product, but permit such a sponsorship in a corporate name." [18] What this means, for example, is that Camel cigarettes can no longer display its logo or colors on a race car, nor can a tennis event be sponsored as the Virginia Slims Legend Tour. However, under the rules, a race car could be sponsored by R.J. Reynolds, minus any product identification, and a tennis tour could be sponsored as the Philip Morris Legends Tour.

The world of motor sports is likely to suffer the most as a result of these regulations. It is estimated that the tobacco industry spent $194 million on sports related sponsorships in 1996.[19] Over the past twenty-five years, R.J. Reynolds (RJR), a leading manufacturer of tobacco products, has spent more than $200 million sponsoring the National Association of Stock Car Auto Racing (NASCAR) series.[20] In the 1996 racing season alone, RJR awarded $4 million in prize money to drivers on NASCAR's Winston Cup Series.[21] Overall, RJR spends over $40 million dollars annually in motor sports programs.[22]

The resolution of the battle between the FDA and the tobacco industry, either by the proposed settlement or the pending appeal in the Fourth Circuit will greatly shape the future of motorsports. The world of auto racing will be looking to fill a huge void left by the loss of millions of dollars in advertising revenue from the tobacco industry. Even worse for the tobacco industry is that they will lose another outlet in which to advertise their product.

This Note examines the constitutionality of the ban on brand-name tobacco sponsorships as it exists under the FDA's final rule as published in the Federal Register. The first part of this Note will provide a historical perspective how the relationship between NASCAR and the tobacco industry has developed over the last twenty-five years. Part Two will discuss the FDA's regulations as published in the Federal Register on August 28, 1996. Part Three will provide a perspective on the development of commercial speech jurisprudence. Parts Four and Five of this Note answer the questions that the FDA did not directly answer—whether the rule violates the First Amendment because the FDA did not differentiate between events attended by adults and those attended by children, and whether there is a solution which would allow brand-name tobacco

---

17. *Id.*

18. FDA Regulations, *supra* note 3, at 44,396. See also *infra* Part II.B for a complete list of the regulations.

19. Bill Koenig, *Auto Racing May Have to Kick Habit*, INDIANAPOLIS STAR, Aug. 24, 1996, at D1.

20. Richard Alm, *Igniting the Opposition*, DALLAS MORNING NEWS, Aug. 24, 1996, at F1.

21. *Id.* R.J. Reynolds, the manufacturer of Winston cigarettes, is a primary sponsor of the elite class of NASCAR racing, the Winston Cup Series, to which it appropriately lent its name.

22. Shav Glick & Jim Peltz, *Up in Smoke?*, L.A. TIMES, Aug. 23, 1996, at C1. This money is spent on prize money, advertising, bonuses and signs at more than 200 race tracks in the United States.

sponsorship in primarily adult events.

I. THE TOBACCO INDUSTRY'S RELATIONSHIP WITH NASCAR

The tobacco industry's relationship with NASCAR began in 1971,[23] shortly after the passage of the Public Health Cigarette Smoking Act of 1969[24] which banned cigarette advertisements on television and radio.[25] The tobacco industry circumvented the problem of not being able to advertise on television by sponsoring auto racing events and racing teams.[26] NASCAR has become the perfect vehicle for the tobacco companies to promote their products.[27]

Once a regional sport, NASCAR has emerged as the fastest growing sport in America.[28] The Winston Cup Series consists of thirty-two races with an average attendance of 171,830 spectators per race.[29] The Winston Cup circuit, traditionally dominant in the South, has expanded its national appeal by holding

---

23. Alm, *supra* note 20, at F1.

24. 15 U.S.C. § 1335 (1994). The Act, as amended, provides: "After January 1, 1971, it shall be unlawful to advertise cigarettes and little cigars on any medium of electronic communication subject to the jurisdiction of the Federal Communications Commission."

25. In *Capital Broadcasting Co. v. Mitchell*, 333 F. Supp. 582 (D.D.C. 1971), six corporations who owned radio stations brought suit seeking to enjoin the enforcement of the ban on radio and television tobacco advertisements. The District Court for the District of Columbia held that the advertising ban did not violate the corporations' First Amendment rights. *Id.* at 584. The court reasoned that the six corporations "ha[d] lost no right to speak - they ha[d] only lost an ability to collect revenue from others for broadcasting their commercial messages." *Id.* The court further held that a rational basis existed for banning cigarette advertisements on television and radio, while allowing them in print ads because "the most persuasive advertising" was on television and radio and that the advertisements were "particularly effective in reaching a very large audience of young people." *Id.* at 585-86.

26. David A. Locke, *Counterspeech as an Alternative to Prohibition: Proposed Federal Regulation of Tobacco Promotion in American Motorsport*, 70 IND. L.J. 217, 218 (1994). Locke stated that with respect to the beneficiaries of tobacco advertising in sports, "the fact that tobacco company names and logos appear on television is incidental to the market competition and bears no relation to the 1971 television ad ban." *Id.* at 223.

27. Alan Friedman, editor of Team Marketing Report, a sports business newsletter, observed that the tobacco industry's nexus between sports is not very significant outside of racing. Alm, *supra* note 20, at F1. Vantage cigarettes sponsors a championship golf tournament and Cambridge cigarettes sponsors bowling. There is a Virginia Slims Legend Tour in women's tennis and a Copenhagen Skoal Pro Rodeo. *Id.* Camel sponsors the professional billiards tour. Glick & Peltz, *supra* note 22, at C1.

28. Scott Moore, *Ladies and Gentlemen, Start Your Televisions*, WASH. POST, May 26, 1996, at Y6.

29. *NASCAR Demographics*, NASCAR PUBLIC RELATIONS MATERIALS. In 1980, 32 Winston Cup races drew 1,555,000 fans for an average attendance of 48,594 spectators per race. Attendance has more than tripled over the last 15 years. In 1995, 5,326,721 fans walked through the turnstiles at 31 Winston Cup events—an increase of 343%. *Id.*

races as far north as New York and New Hampshire and as far west as California and Arizona.[30] New race tracks in Texas, Las Vegas, and southern California opened their gates to NASCAR in 1997.[31] NASCAR's marketing strategy includes incorporating its fans into the growth of the sport. In February of 1996, NASCAR opened Daytona USA, at the Daytona International Speedway in Daytona Beach, Florida. The attraction allows fans to drive in a simulated race or be a part of a mock pit-stop.[32] NASCAR has opened a line of NASCAR licensed stores and theme restaurants.[33] As a result of its growth, merchandising revenues from NASCAR licensed products are at an all-time high. Merchandise sales have increased from $60 million in 1990 to $700 million in 1995,[34] and are projected to have surpassed the $1 billion mark in 1996.[35]

The growth of the sport has resulted in increased media exposure. Every NASCAR event is televised.[36] In 1995, television viewership exceeded 120 million, an increase of twenty-three percent over 1994.[37] Over the past five years, NASCAR ratings have increased[38] as broadcast and cable networks have increased the amount of time they devoted to race programming.[39] In May of 1996, a die-hard race fan could have watched more than one-hundred hours of live race programming.[40]

NASCAR is appealing to sponsors for several reasons. For the tobacco companies, the most obvious reason is that sponsoring race events and racing teams is a vehicle to promote tobacco products on television despite the

---

30. Chris Roush, *Red Necks, White Socks, and Blue-Chip Sponsors*, BUS. WK., Aug. 15, 1994, at 74.

31. Alm, *supra* note 20, at F1.

32. *Marketing of NASCAR Continues to Expand*, CHARLESTON DAILY MAIL, July 29, 1996, at B3.

33. *Id.*

34. Rick Maloney, *Once Around NASCAR Store Brings the Track to the Fan*, BUSINESS FIRST OF BUFFALO, July 29, 1996 (page references are not available). In 1994, Dale Earnhardt, a seven-time champion on the Winston Cup Circuit, sold over $900,000 worth of merchandise on QVC in just over two hours. Roush, *supra* note 30, at 74.

35. Roush, *supra* note 30, at 74. Steven M. Bornstein, ESPN, Inc.'s chief executive, said, "The NASCAR guys are the brightest marketers I've known. They understand who they're trying to appeal to, and they've developed some of the brightest racing stars." *Id.*

36. Moore, *supra* note 28, at Y6.

37. Glick & Peltz, *supra* note 22 , at C1.

38. Moore, *supra* note 28, at Y6.

39. *Id.* In 1996, The Nashville Network (TNN) scheduled 800 hours of motor sports; ESPN televised sixteen NASCAR Winston Cup races; ESPN2 planned 700 hours of motor sports programming; TBS's Winston Cup ratings increased 23% in 1995 and broadcasted three Winston Cup races in 1996; CBS scheduled 30 hours for 13 NASCAR events, including the Daytona 500, whose ratings increased to 18.2%, reaching 8.82 million homes; and, ABC scheduled 60 hours of racing, including the Indianapolis 500. *Id.*

40. *Id.*

advertising ban.[41] According to Joyce Julius & Associates, a company that studies media time, Winston cigarettes received television exposure worth an estimated $944,000 during CBS's broadcast of the 1996 Daytona 500.[42] Furthermore, sports sponsorship is cost effective. Five million dollars will sponsor a top-of-the-line Winston Cup racing team for the entire season.[43] Compared to the cost of a thirty second ad on national prime-time television, this is a bargain.[44]

There are other substantial reasons, which few other sports can boast, why a company marketing its product would want to use NASCAR as a marketing tool. A primary reason is the demographics of the more than five million people who attend races each year. Thirty-eight percent of the NASCAR audience is female, while the average annual household income is $39,280.[45] Moreover,

---

41. *See supra* notes 24-25 and accompanying text.

42. Alm, *supra* note 20, at F1. A study conducted by Joyce Julius & Associates revealed that during the 1993 racing season, "tobacco company sponsors' 'names, product(s), or clearly recognizable slogan(s)' received televised mention no fewer than 3675 times . . . ." Locke, *supra* note 26, at 223. The report estimated the total value of tobacco industry's television exposure to be over $31 million. *Id.* at 223 n.37. The value of the sponsor's television exposure was calculated by adding the time of "in-focus" exposure of the product ("in-focus" exposure includes "car identity, uniforms, helmets/hats, shirts, billboards, signs, retaining walls, pit identification, starter's stand, television screen graphics, car transporters, flags, banners, message boards and scoreboards. *Id.* at 223 n.34) and the number of verbal references made to the product. Each verbal reference was rated at 10 seconds. The dollar value of the total exposure was calculated by using the advertising rate for a thirty second commercial during the race broadcast. *Id.* at 223 n.37. During the 1993 NASCAR season, tobacco companies received 16 hours, 5 minutes, 27 seconds of "in-focus" exposure, 1918 sponsor mentions and $31,379,755 in advertising value. *Id.* (citing Joyce Julius & Associates, 9 SPONSOR'S REP. NO. 32, 1993 NASCAR/WINSTON CUP YEAR END REPORT, at 2-5 (1993)).

43. Roush, *supra* note at 30, at 74.

44. Randall H. Stoner, *200 MPH Cigarette Ads: A Comparison of International Restrictions on Tobacco Sports Sponsorship*, 15 HASTINGS INT'L & COMP. L. REV. 639, 643-44 (1992).

> To advertise for 30 seconds on NBC's Bill Cosby Show costs 250,000 dollars. Now a million bucks spent in NASCAR will get you a middle-of-the-pack Winston Cup team, a team that will generate literally hours of exposure on national television for thirty races throughout the year. And it will do so for the cost of two minutes of national advertising.

Id. (quoting Dutch Mandel, *Autopower in the 90's; Advertising & Marketing the Persuaders*, AUTOMOTIVE NEWS, Nov. 29, 1989, at 120).

45. *Marketing of NASCAR Continues to Expand*, *supra* note 32, at B3. *See also NASCAR Demographics*, NASCAR PUBLIC RELATIONS MATERIALS, (Sources: Simmons Market Research Bureau, Inc. and Performance Research). The following is a complete demographic breakdown of the NASCAR audience as provided by NASCAR:

NASCAR fans are extremely brand-loyal.[46] Over seventy percent of NASCAR fans said that they would "almost always" or "frequently" buy a product involved in NASCAR.[47] Compared to baseball at fifty-eight percent, tennis at fifty-two percent, and golf at forty-seven percent, NASCAR's product-loyal fans "put[] a gleam in every marketer's eye."[48]

While cigarettes, beer and auto parts are still a mainstay in racing sponsorship, NASCAR's sponsorship list has diversified, attracting consumer products like Kodak film, Kellogg's cereal, Maxwell House coffee, Tide laundry detergent, McDonald's, Pepsi, and Gatorade.[49] Prodigy, an on-line computer service, and cable television networks, like the Cartoon Channel and QVC, have joined the frenzy to be a part of the sport.[50]

Corporate sponsorship of NASCAR is expected to grow to $441 million in

---

GENDER	
Female	38%
Male	62%

INCOME	
less than $10,000	7%
$10,000-19,999	14%
$20,000-29,999	17%
$30,000-39,999	19%
$40,000-49,999	14%
over $50,000	29%

MARITAL STATUS	
Married	64%
Single	22%
Divorced/Widowed	14%

EMPLOYMENT	
Full-time	72%
Part-time	10%
Retired/Unemployed	18%

OCCUPATION	
Professional/Managerial	27%
Technical/Clerical/Sales	21%
Craft Precision	13%
Unskilled Labor	10%
Other	29%

RESIDENCE	
Rent	28%
Own	72%

EDUCATION	
Some High School	12%
Graduated High School	88%
Some College/Graduated College	38%

AGE	
Under 18	3%
18-24 yrs. old	15%
25-34 yrs. old	29%
35-44 yrs old.	25%
45-54 yrs. old	16%
55 and over	12%

46. Roush, *supra* note 30, at 74.

47. *Id.*

48. *Id.*

49. *1996 Official Family of Sponsors*, NASCAR PUBLIC RELATIONS MATERIALS.

50. *Marketing of NASCAR Continues to Expand*, *supra* note 32, at B3. The Cartoon Network sponsored race car sports the image of Fred Flintstone. According to Christopher Jones, an official with Turner Broadcasting (Turner owns the Cartoon Network), NASCAR fans are the type of fans that they want to reach. "[NASCAR] appeals to the masses while other sports appeal to the elites. Our characters are working-class characters. Fred [Flintstone] carries a lunch pail to work everyday." *Id.*

1997, up from $405 million in 1996.[51] By comparison, all other stadium sports combined generated only $365 million in corporate support in 1996.[52] Despite the long line of companies, "such as Proctor & Gamble, DuPont, and McDonald's [who] vie for the opportunity to splash $4 million apiece on these 200-mph advertising vehicles,"[53] if the tobacco companies are prohibited from pouring cash into NASCAR's coffers, the sport will definitely feel an impact.[54] While there is a consensus that racing would feel the impact of a tobacco advertising ban, the severity is unknown.[55] Moreover, the FDA's regulations are one more strike against the tobacco industry and the social stigma attached to smoking. Undoubtedly, the real loser from the FDA's regulations, is not NASCAR, but the tobacco industry.

II. THE FOOD AND DRUG ADMINISTRATION'S REGULATION OF TOBACCO

A. Brief Historical Overview of Tobacco Regulation

Historically, the FDA has taken the position that tobacco regulation is outside of its domain.[56] The FDA's predecessor, the Bureau of Chemistry in the

---

51. Lee Walczak et al., *Speed Sells*, BUS. WK., Aug. 11, 1997, at 86.

52. *Id.*

53. Alex Taylor III, *Can NASCAR Run in Bigger Circles?*, FORTUNE MAG., Feb. 5, 1996, at 38.

54. Locke, supra note 26, at 224.

55. *Id.* Hardy Smith, the executive director of the National Motorsports Council, an organization which was formed to lobby on issues which affect the sport of racing, said that under a tobacco advertising ban "[c]ompetition would suffer, and races might be moved from free TV to pay-per view." *Id.* at 224 n.46 (quoting Liz Clarke, *Will Sponsorship Go Up in Smoke?*, INDIANAPOLIS STAR, Aug. 4, 1994, at E26). *But see* Jeff Jensen, *Non-Tobacco Sponsors Could Fill Motor Void*, ADVERTISING AGE, Sept. 2, 1996, at 34. ("Would it be catastrophic for NASCAR if Winston (tobacco sponsorship) went away? No, it wouldn't. But that doesn't mean we aren't going to fight.") (quoting John Story, director of public relations for the Daytona International Speedway).

On June 20, 1997, just after the proposed settlement was announced, NASCAR issued an official statement on the recent developments between the tobacco industry and the government.

> For nearly 30 years we have had a mutually beneficial relationship with R.J. Reynolds and its Winston brand. We have not had an opportunity to review the proposed agreement so it would be premature to speculate on what effect this will have on motorsports. While the settlement has been announced, it must still face Congressional as well as Presidential review while also facing litigation that has already been filed. NASCAR racing has been in existence for nearly 50 years, long before tobacco companies became actively involved in the sport. With monumental growth we have experienced in recent years, and the anticipated continued growth of motorsports, we will continue to aggressively promote the sport.

NASCAR Tobacco Statement, June 20, 1997 <http:\\www.nascar.com>.

56. U.S. Dep't. of Health and Human Servs., *FDA's Proposed Regulation of the Sale and*

Department of Agriculture, announced in 1914 that tobacco, which had not been labeled for "medicinal purposes," was outside the scope of the Federal Food and Drug Act of 1906.[57] In 1938, the Federal Food, Drug, and Cosmetic Act (FDCA) was passed.[58] The FDCA is an enabling statute which provides the FDA with the authority to promulgate regulations.[59] Under the FDCA, the FDA has the authority to regulate "food," "drug," "device," and "cosmetics" as defined by the statute.[60] Since its enactment in 1938, tobacco has not been interpreted to fit into any one of these definitions, nor has Congress passed legislation which would give the FDA jurisdiction.[61] However, on August 11, 1995, the FDA, under existing law, issued a proposal which conferred upon themselves the authority to regulate tobacco.[62] On August 28, 1996, the FDA issued a final rule, *Regulations Restricting the Sale and Distribution of Cigarettes and Smokeless Tobacco to Protect Children and Adolescents*, which is to become fully effective by February 28, 1998.[63]

Promotion of Tobacco Products to Minors, Public Health Rep. 1996; 111, 280-85.

57. Whatley, *supra* note 2, at 122.

58. Federal Food, Drug, and Cosmetic Act, Pub. L. No. 75-717, 52 Stat. 1040 (1938) (codified as amended at 21 U.S.C. § 301-395 (1994)).

59. Boeckman, *supra* not 2, at 997 n.23. An administrative agency is built upon the foundation of an enabling statute, "so that acts exceeding the scope of the statute are invalid." *Id.*

60. 21 U.S.C. §§ 321(f) -(i) (1994).

61. Over the years, bills have been introduced in Congress to grant the FDA regulatory authority over tobacco. However, none of these bills, which would have given the FDA jurisdiction via the legislature, have passed. Whatley, *supra* note 2, at 122-25. In 1956, the House rejected a bill that would have amended the FDCA, giving the FDA jurisdiction over cigarettes. *Id.* at 122 (H.R. 11280, 84th Cong., 2nd Sess. (1956)). In 1963, bills were rejected by the House and Senate that would have put smoking products under the FDA's jurisdiction. *Id.* (H.R. 5973, 88th Cong., 1st Sess. (1963); S. 1682, 88th Cong., 1st Sess. (1963)). In 1964, legislation was introduced once again to give the FDA jurisdiction over tobacco. Instead, Congress passed the Federal Cigarette Labeling and Advertising Act of 1965 to deal with the issue of advertising and labeling with respect to smoking and health. *Id.* at 122-23. In 1972, FDA Commissioner Charles Edwards said at Congressional hearings that "the regulation of cigarettes is to be the domain of Congress . . . and labeling or banning cigarettes is a step that can be taken only by Congress." *Id.* at 123 (quoting Hearings Before the Consumer Subcomm. of the Senate Comm. on Commerce on S. 1454, 92nd Cong., 2d Sess. (1972)). In 1986, Congress passed the Comprehensive Smokeless Tobacco Health Education Act (CSTHEA) to provide a federal regulatory scheme for the regulation of smokeless tobacco products. *Id.* at 124 (15 U.S.C. §§ 4401-4408 (1994)). The CSTHEA does not grant any jurisdiction to the FDA. *Id.* In 1987, 1989, 1992, and 1993, bills were introduced in Congress attempting to place tobacco products within the reach of the FDA; none of them passed. *Id.* at 124-25 (H.R. 3294, 100th Cong., 1st Sess. (1987); H.R. 1494 and S. 769, 101st Cong., 1st Sess. (1989); H.R. 4350 and S. 2298, 103d Cong., 1st Sess. (1992); H.R. 2147 and S. 672, 104th Cong., 1st Sess. (1993).

62. Proposed Regulations, *supra* note 16.

63. FDA Regulations, *supra* note 3, at 44,396. The FDA published a 924 page document in two volumes in the August 28, 1996 Federal Register. Book One sets forth the regulation and

B. The FDA Rule

The purpose behind the FDA Rule is to "restrict[] the sale and distribution of cigarettes and smokeless tobacco to children and adolescents" which will "prevent future generations of Americans from becoming addicted to them [while allowing] the continued marketing of these products."[64] The rule will:

> (1) prohibit the sale of nicotine-containing cigarettes and smokeless tobacco to individuals under the age of eighteen;
>
> (2) require manufacturers, distributors, and retailers to comply with certain conditions regarding the sale and distribution of these products;
>
> (3) require a retailer to verify a purchaser's age by photographic identification;
>
> (4) prohibit all free samples and prohibit the sale of these products through vending machines and self-service displays except in facilities where individuals under the age of eighteen are not present or permitted at any time;
>
> (5) limit the advertising and labeling to which children and adolescents are exposed to a black-and-white, text-only format;
>
> (6) prohibit the sale or distribution of brand-identified promotional nontobacco items such as hats and tee shirts;
>
> (7) prohibit the sponsorship of sporting and other events, teams, and entries in a brand name of a tobacco product, but permit such sponsorship in a corporate name; and
>
> (8) require manufacturers to provide intended use information on all cigarette and smokeless tobacco product labels and in cigarette advertising.[65]

The FDA asserts that the regulations "will address the serious public health problems caused by cigarettes and smokeless tobacco products;" that they "will reduce children's and adolescents' easy access" to tobacco products; and that they will "significantly decrease the amount of positive imagery that makes [tobacco] products so appealing to that age group."[66]

---

the FDA's justifications and need for tobacco regulation. Book Two consists of an "Annex" which covers tobacco research and regulation in general.

64. *Id.* at 44,397.

65. *Id.* at 44,396.

66. *Id.*

The FDA predicated their assertion of jurisdiction over tobacco products under the FDCA, by classifying cigarettes and smokeless tobacco as "delivery devices" for nicotine.[67] The reversal in the FDA's policy did not come from any change in the law, rather it came from evidence that the tobacco companies control and manipulate the levels of nicotine in tobacco products.[68] With this ammunition, the FDA classified cigarettes and smokeless tobacco as "combination products consisting of a drug[69] (nicotine) and device components intended to deliver nicotine to the body."[70] The FDA, having the discretion to choose which authority to apply in the regulation of the combination products, "determined that tobacco products are most appropriately regulated under the device provisions of the act."[71]

C. The FDA's Justification for Imposing a Brand-Name Sponsorship Ban

Under the FDA's sponsorship ban, "no manufacturer, distributor, or retailer shall sponsor or cause to be sponsored any athletic, musical, artistic or other social or cultural event, in the brand name, logo, motto, selling message, recognizable color or pattern of colors, or any other indicia of a product identification similar or identical to those used for tobacco or smokeless tobacco products."[72]

---

67. *Id.* at 44,400. A "device" is defined as "instruments, apparatus, and contrivances, including their components, parts, and accessories, intended . . . to affect the structure or any function of the body of man." 21 U.S.C. § 321(h)(3) (1994).

68. U.S. Dep't. of Health and Human Servs., *supra* note 55, at 280. On December 8, 1996, CBS ran a story featuring David Kessler, the commissioner of the FDA who announced in November, 1996 that he was stepping down from his position as head of the FDA. Through an extensive investigation, Kessler uncovered that tobacco manufacturers were manipulating the nicotine levels in tobacco through genetic and chemical engineering. Kessler discovered a patent filed by Brown & Williamson, a major tobacco manufacturer, that was registered in Brazil, written in Portuguese, and filed in the Netherlands. The patent was for tobacco plants with a six percent nicotine content, twice the level of the highest flue-cured in the U.S. The tobacco plants were grown in Brazil and flown back into the U.S. for manufacture. Kessler also uncovered a manual from Brown & Williamson which described the effect of adding ammonia and other chemicals to tobacco to enhance the effects of nicotine. *60 Minutes: How He Won the War* (CBS television broadcast, Dec. 8, 1996) *available in* WESTLAW 1996 WL 8065061.

69. A drug is defined as an article other than food which is "intended to affect the structure or any function of the body of man." 21 U.S.C. § 321(g)(1)(c) (1994). The FDA determined that cigarettes and smokeless tobacco products are "'drugs' which produce[] significant pharmacological effects in consumers, including satisfaction of addiction, stimulation, sedation, and weight control." FDA Regulations, *supra* note 3, at 44397.

70. *Id.* at 44,400.

71. *Id.* For articles examining the statutory interpretation of the FDCA as applied to tobacco regulation and separation of powers issues raised by the FDA's regulations, see Harder, *supra* note 11 and Boeckman, *supra* note 2.

72. FDA Regulations, *supra* note 3, at 44,527 (codified at 21 C.F.R. 897.34(c) (1996)).

While brand-name sponsorship is prohibited under this rule, a tobacco company may still sponsor an event, team or entry as a corporate sponsor. However, colors and patterns of colors which would be recognizable with a brand name product are not allowed. The product names and logos of cigarettes and smokeless tobacco, such as Winston, Camel, Marlboro, and Skoal, will not be permitted to be plastered on every surface imaginable around the race track. Winston Cup tee shirts and caps, a sight common at every race track, will no longer be permitted to be sold or distributed. Thus, the incentive that a tobacco company will have for sponsoring an event—product recognition—is eliminated.

The FDA asserts that brand-name sponsorship by tobacco companies "associates tobacco use with exciting, glamorous, or fun events such as car racing and rodeos . . . [and] creates a 'friendly familiarity' between tobacco and sports enthusiasts, many of whom are children and adolescents."[73] The intentions of the regulations are to "reduce the 'friendly familiarity'."[74] The agency contends that the regulations are drafted narrowly to protect children and adolescents, while "recogniz[ing] the importance of corporate sponsorship in engendering goodwill . . . [by] providing support to sports, the arts, and music."[75]

While the sponsorship ban would impose an across-the-board ban on all types of tobacco sponsorships, the FDA, in justifying its rule, was clearly concerned about motorsport sponsorships, and NASCAR in particular. Thus, the majority of the regulation concerning the sponsorship ban was directed towards demonstrating the effect of motorsport sponsorship on children under eighteen years old. [76]

The FDA found that the number of children affected by motorsport sponsorship was substantial. The FDA relied on data which showed that 354 motorsport events had a total viewing audience of 915 million people, of which 64.05 million were under the age of eighteen.[77] This averaged to 180,806 underage television viewers per event.[78] The FDA also noted that the number of children in actual attendance at racing events "may be growing."[79]

The FDA claimed that the effect that tobacco sponsorship had on children and adolescents is "enormous."[80] The FDA stated that sponsorship creates an "attractive and exciting image" of tobacco products "that can serve as a 'badge' or identification," and because of the prolonged exposure of a product in a sponsored event, "an impression of prevalence and normalcy about tobacco use" is created.[81] They further asserted that children "will repeatedly see and begin

---

73. *Id.*

74. *Id.*

75. *Id.* at 44,536.

76. *See generally* FDA Regulations, *supra* note 3, at 44,527-44,536.

77. *Id.* at 44,528.

78. *Id.*

79. *Id.* NASCAR reports that only three percent of spectators at NASCAR Winston Cup Races are under eighteen years-old. NASCAR PUBLIC RELATIONS MATERIALS, *supra* note 29.

80. FDA Regulations, *supra* note 3, at 44,528.

81. *Id.* at 44,529.

to associate the event, which they are enjoying, with the imagery and appeal of the product."[82]

According to the FDA, children and adolescents "are still forming attitudes and beliefs about tobacco use" and see smoking "as a coping mechanism, a gauge of maturity, a way to enter a new peer group, or as a means to display independence."[83] The rule is intended to break the link between tobacco brand-sponsored events and images and use of tobacco by young people."[84]

The FDA received criticism that the rule is overly broad and violates the First Amendment.[85] The rule prohibits brand-name tobacco sponsorships at all events and does "not attempt to differentiate between those events that attract children and adolescents and those that attract adults."[86] The FDA did not directly provide a response to this question. Instead, they skirted around the issue by stating that children who attend events are "directly and unavoidably confronted with messages from the sponsoring product" and that viewers on television are made aware of particular brands.[87] Thus, considering these factors, a sponsorship ban will effectively "limit the influences on children . . . and thus, protect their health."[88] The FDA further stated that they were "not aware of any way to limit the restriction to events that are attended by young people."[89]

III. THE COMMERCIAL SPEECH DOCTRINE

A. Historical Perspective on Commercial Speech

Commercial speech is a doctrine in search of a theory. From the time it showed up on the doorstep in 1942 (and carrying quite a lot of baggage at that), commercial speech has been the poor relation of the

---

82. *Id.*
83. *Id.* at 44,530.
84. *Id.*
85. *Id.* at 44,533.
86. *Id.*
87. *Id.*
88. *Id.*
89. *Id.* at 44,534. The FDA seems to be inconsistent in this statement. The FDA has developed a standard which would differentiate among adults and adolescents. In restricting printed advertisements to black and white text only, the FDA proposed a "print media standard" where the "tombstone" ads, to which the black and white text ads are referred, are only applicable to magazines where children comprise at least 15% or two million of the readers. FDA Regulations, *supra* note 3, at 44,513. However, the FDA considered and rejected this proposal for sponsorships. They reasoned that the types of exposure are "dramatically different" because the time spent viewing an advertisement is nine seconds, while during an event, the viewer or spectator is "unavoidably bombarded with posters, signs, hats . . . linked with a fun, exciting, or glamorous event that they enjoy for a prolonged period of time." *Id.* at 44,529. *See infra* discussion Part IV.D.

First Amendment family.[90]

Commercial speech has been defined by the Court as speech that does "no more than propose a commercial transaction."[91] In 1942, the Supreme Court was first confronted with whether commercial speech was afforded First Amendment protection in *Valentine v. Christensen*.[92] The Court held that it was not.[93] The Supreme Court's opinion consisted of only two paragraphs, which quickly disposed of the matter by stating that "the Constitution imposes no [First Amendment] restraint on government as respects purely commercial advertising."[94] Despite criticism[95] and apparent inconsistencies,[96] the Court

---

90. RICHARD T. KAPLAR, THE MEDIA INSTITUTE, ADVERTISING RIGHTS THE NEGLECTED FREEDOM: TOWARD A NEW DOCTRINE OF COMMERCIAL SPEECH 35 (1991).

91. Pittsburgh Press Co. v. Pittsburgh Comm'n on Human Relations, 413 U.S. 376, 385 (1973).

92. 316 U.S. 52 (1942). In *Valentine*, Christensen owned a submarine exhibit in New York City. Christensen prepared handbill advertisements soliciting visitors to his submarine for a stated admission. Attempting to distribute the handbills on the streets of New York, Christensen was told that this activity violated the Sanitary Code, which prohibited the distribution of commercial and business advertisements in the streets. Subsequently, Christensen made two-sided handbills. One side contained an advertisement for his submarine, the other contained a protest with no commercial advertisement. The police restrained Christensen from distributing the new handbills. The Court was unpersuaded by the fact that Christensen printed a protest on one side of the handbill. The Court concluded that the affixing of the protest to the commercial handbill "was with the intent, and for the purpose, of evading the prohibition of the ordinance." *Id.* at 55.

93. *Id.*

94. *Id.* at 54. The Court was only partially correct in this statement. The First Amendment is silent with respect to commercial advertising and its relationship to the First Amendment. Thus, this issue is not as black and white as the Court makes it seem. *See* KAPLAR, *supra* note 89, at 36-46 (discussion of whether the founders intended to protect commercial speech); *see also* Locke, *supra* note 26, at 240 ("[w]hile it is reasonable to assume that the Founding fathers did not intend the First Amendment to shield speech which inflicts injury or is calculated to incite immediate breaches of the peace, it is less clear that the Amendment . . . allow[s] courts to weigh the value of certain categories [other than 'fighting words']. In the case of tobacco sports advertising and promotion, the speech itself does not trigger the expected harm . . . of the 'fighting words' doctrine; rather, [it is] only indirectly related to the interest of the government in preventing illness and disease.").

95. Justice Douglas said in 1959 that the *Valentine* "ruling was causal, almost offhand. And it has not survived reflection." KAPLAR, *supra* note 89, at 18 (quoting Cammarano v. U.S., 358 U.S. 498, 514 (1959) (Douglas, J., concurring)).

96. In *New York Times v. Sullivan*, 376 U.S. 254 (1964), the *Times* published an advertisement that allegedly injured the reputations of certain public officials. The Court held that the advertisement was afforded First Amendment protection because it communicated "claimed abuses and sought financial support on behalf of a movement whose existence and objectives are matters of the highest public concern." *Id.* at 266. Bradford W. Scharlott criticized the Court's decision, stating that in principle the *Times* case was no different than *Valentine*. Scharlott stated

followed this doctrine for thirty-three years, holding that commercial speech received no protection under the First Amendment.[97] However, in 1975, the Court modified their position by reformulating a test for determining whether commercial speech was within the ambit of the First Amendment. The Court in *Bigelow v. Virginia*[98] held that *Valentine* "obviously does not support any sweeping proposition that advertising is unprotected *per se*;"[99] rather, the constitutionality of commercial speech should be "assess[ed] [by] the First Amendment interest at stake and weighing it against the public interest allegedly served by the regulation."[100]

In 1976, the commercial speech doctrine reached its "high-water mark"[101] in *Virginia State Board of Pharmacy v. Virginia Citizens Consumer Council, Inc.*[102] The Court strengthened its revised position on commercial speech, holding that just because the "the advertiser's interest is a purely economic one," that does not "disqualif[y] him from protection under the First Amendment."[103] In *Virginia Board of Pharmacy*, Virginia consumers challenged a state statute that prohibited the advertisement of prices for prescription drugs.[104] A U.S. District Court struck down the statute,[105] and Justice Blackmun writing for the Supreme Court upheld the district court's ruling.[106]

Virginia Board of Pharmacy was significant because the issues dealt squarely with a commercial transaction and did not contain any other underlying concerns. As the Court stated:

> Our pharmacist does not wish to editorialize on any subject, cultural, philosophical, or litical. He does not wish to report any particularly newsworthy fact, or to make generalized observations even about commercial matters. The "idea" he wishes to communicate is simply

---

that "[i]n both cases, the advertisers had a commercial goal, but also expressed grievances." Bradford W. Scharlott, *The First Amendment Protection of Advertising in the Mass Media, in* ADVERTISING AND COMMERCIAL SPEECH, 4 (Hon. Theodore R. Kupferman ed., 1990). Yet, the outcome was completely different. Scharlott characterized the *Valentine* test as "deficient" because it produced "an all-or-nothing" outcome—the advertisement either received full First Amendment protection or it received none. *Id.*

97. *See, e.g.*, Williamson v. Lee Optical, Inc., 348 U.S. 483 (1955).

98. 421 U.S. 809 (1975).

99. *Id.* at 820.

100. *Id.* at 826. In *Bigelow*, the Court held that a statute which prohibited all abortion advertisements was unconstitutional. The Court reasoned that the advertisements "did more than simply propose a commercial transaction. [The advertisements] contained factual material of clear 'public interest.'" *Id.* at 822.

101. KAPLAR, *supra* note 89, at 22.

102. 425 U.S. 748 (1976).

103. *Id.* at 762.

104. *Id.* at 749-50.

105. *Id.* at 750.

106. *Id.* at 773.

> this: "I will sell you the X prescription drug at the Y price." Our question, then, is whether this communication is wholly outside the protection of the First Amendment.[107]

In holding that a pure commercial transaction is afforded First Amendment protection, the Court reasoned that "speech does not lose its First Amendment protection because money is spent to project it."[108] The Court broke "new ground" by holding that "[if] there is a right to advertise, there is a reciprocal right to receive the advertising."[109] Despite "however tasteless and excessive" advertising may be, the Court said that it is "nonetheless dissemination of information as to who is producing and selling what product, for what reason, and at what price. . . . To this end, the free flow of commercial information is indispensable."[110]

Despite what appeared to be victory for the constitutional protection of commercial speech, in what appears almost as an afterthought, in a footnote near the end of the opinion, the Court "held that commercial speech is entitled to only a second-class level of First Amendment protection."[111] The Court stated that even though commercial speech is afforded protection under the First Amendment, "[t]here [are] some commonsense differences between speech that does 'no more than propose a commercial transaction'" and other types of speech which "suggest that a different degree of protection is necessary."[112] Thus, commercial speech enjoyed " a limited measure of protection, commensurate with its subordinate position in the scale of First Amendment values."[113]

Although *Virginia Board of Pharmacy* established that commercial speech was afforded constitutional protection, in 1980 the Court formulated a four part test in *Central Hudson Gas & Electric v. Public Service Commission of New York*[114] to determine if, despite First Amendment protection, commercial speech could be regulated.

B. Central Hudson

In *Central Hudson* the Court established the framework for the modern analysis of commercial speech, promulgating a four part test to determine

---

107. *Id.* at 761.

108. *Id.*

109. KAPLAR, *supra* note 89, at 21 (quoting *Virginia Board of Pharmacy*, 425 U.S. at 757).

110. *Virginia Board of Pharmacy*, 425 U.S. at 765.

111. Scharlott, *supra* note 95, at 6.

112. *Virginia Board of Pharmacy*, 425 U.S. at 772 n.24. *See* KAPLAR, *supra* note 89.

113. Florida Bar v. Went For It, Inc., 515 U.S. 618, 623 (1995). In *Went For It*, lawyers brought action to challenge the constitutionality of the Florida Bar rules which prohibit lawyers from using direct mail to solicit personal injury or wrongful death clients within thirty days of the accident. The statute was upheld as lawyer solicitation was found to be commercial speech subject to intermediate scrutiny. *Id.*

114. 447 U.S. 557 (1980).

whether the speech should be afforded constitutional protection.[115] The threshold inquiry is whether (1) the communication is misleading or unlawful.[116] If the answer is yes, then the inquiry stops and the regulation is upheld.[117] However, if the activity is lawful and the communication is not misleading, then (2) the government must assert a "substantial interest;" [118] (3) the regulation must "directly advance" the government's interest;[119] and (4) the regulation must be "no more extensive than necessary" to further the government's interest.[120]

In 1973, the New York Public Service Commission ordered all electric utilities in New York state "to cease all advertising that "promot[ed] the use of electricity."[121] Due to the fuel shortage, the commission was concerned that there was not enough fuel to meet customers' demands during the 1973-74 winter. After the fuel shortage ended, the commission extended the prohibitions on advertising in a Policy Statement issued February 25, 1977. The Policy Statement divided advertising into two categories: (1) "promotional—advertising intended to stimulate the purchase of utility services," and (2) "institutional and informational, a broad category inclusive of all advertising not clearly intended to promote sales."[122] Institutional and informational advertising was allowed; however, all promotional advertising was banned. The Commission reasoned that promotional advertising was "contrary to the national policy of conserving energy."[123]

Central Hudson challenged the Commission's order, arguing that the advertising ban "restrained commercial speech in violation of the First and Fourteenth Amendments."[124] The order was upheld at the trial court and at the intermediate appellate level.[125] The New York Court of Appeals affirmed both courts, holding that "the governmental interest in the prohibition outweighed the limited constitutional value of the commercial speech at issue."[126] The Supreme

---

115. Richard T. Kaplar likened the *Central Hudson* test to a "patchwork quilt" where the Court's previous decisions were finally piece-mealed together into one opinion. Kaplar notes that the *Central Hudson* opinion "bore a stronger resemblance to *Bigelow*, with its emphasis on the balancing of competing values, than to *Virginia Pharmacy Board's* reliance of the primacy of the First Amendment." KAPLAR, *supra* note 89, at 23.

116. *Central Hudson*, 447 U.S. at 563. The Court stated that "there can be no constitutional objection to the suppression of commercial messages that do not accurately inform the public" or to "commercial speech related to illegal activity." *Id.* at 563-64.

117. *Id.* at 563.

118. *Id.* at 564.

119. *Id.*

120. *Id.* at 569-70.

121. *Id.* at 558.

122. *Id.* at 559.

123. *Id.*

124. *Id.* at 560.

125. *Id.* at 560-61.

126. *Id.* at 561.

Court granted *certiorari* and reversed.[127]

Initially, the Court observed that it has previously "rejected the 'highly paternalistic' view that government has complete power to suppress or regulate commercial speech" because "[p]eople will perceive their own best interests if only they are well enough informed, and . . . the best means to that end is to open the channels of communication rather than to close them . . . ."[128] However, the Court noted the "commonsense" distinction between commercial speech and other types of speech and held that the Constitution "accords a lesser protection to commercial speech than to other constitutionally guaranteed expression."[129]

In applying the four part test, the Court found that the advertising was commercial speech that was protected by the First Amendment, thus, satisfying the first prong.[130] The Court also held that the Commission asserted a substantial state interest and that the regulation directly advanced the state interest.[131] However, with respect to the fourth prong, the Court held that the advertising ban was more extensive than necessary; thus, the regulation failed.[132]

C. "In Search of a Theory"

The commercial speech doctrine is in a state of uncertainty. While the *Central Hudson* test has not been overruled, the Court has struggled to find a consistent application of the test.[133] In all but expressly overruling a 1986

---

127. *Id.*

128. *Id.* at 562 (citations omitted).

129. *Id.* at 562-63.

130. *Id.* at 567. The Commission did not claim that the communication was unlawful or misleading. Despite arguments that the speech was not protected because Central Hudson held a monopoly in electric service, the Court held otherwise because "a monopoly enterprise legitimately may wish to inform the public that it has developed new services or terms of doing business." *Id.*

131. The Commission asserted two state interests: (1) the need for energy conservation, and (2) an equitable rate structure for utility costs. *Id.* at 568-69. The Court found that these interests were substantial and that the state's interest in energy conservation was "directly advanced" by the restrictions. *Id.* at 569.

132. *Id.* at 570. The Court reasoned that the "commission failed to show that "its interest in conservation cannot be protected adequately by more limited regulation of appellant's commercial expression." *Id.* The Court suggested that the Commission could "require that the advertisements include information about the relative efficiency and expense of the offered service . . . ." *Id.* at 571. The Court further explained what was required for the fourth prong to be satisfied in a 1989 decision.

> What our decisions require is a "'fit' between the legislature's ends and the means chosen to accomplish those ends,"—a fit that is not necessarily perfect, but reasonable; . . . that employs not necessarily the least restrictive means but . . . a means narrowly tailored to achieve the desired objective.

Board of Trustees v. Fox, 492 U.S. 469, 480 (1989) (citations omitted).

133. The Court has stated that all that was required with respect to the third and fourth prongs was a "reasonable fit." However, in subsequent cases, the Court warned that "the restrictions must

decision,[134] the Supreme Court has changed its position on commercial speech and appears ready to depart from *Central Hudson* and afford commercial speech more protection.

1. Posadas de Puerto Rico Associates v. Tourism Company.[135]—In *Posadas*, the Court seemed to defer to the judgment of the legislature and afforded less constitutional protection to advertisements of "vice" activities. In *Posadas*, the operator of a Puerto Rican casino challenged a statute which restricted advertising of casino gambling aimed at Puerto Rican residents, but permitted advertising directed towards tourists. The Court applied the *Central Hudson* test and affirmed the Supreme Court of Puerto Rico's decision holding the statute facially constitutional.[136] The Court held that the commercial speech was neither unlawful nor misleading and that Puerto Rico had a substantial interest in the "health, safety, and welfare of its citizens . . . ."[137] However, with respect to the third and fourth prongs of the *Central Hudson* test, the Court deferred to the judgment of the legislature in holding that they were satisfied.[138] The effect of this decision is significant because the two most difficult prongs to satisfy were substantially weakened.[139]

---

be 'narrowly tailored,' cannot 'burden substantially more speech than necessary, and that the existence of numerous and obvious less-burdensome non-speech alternatives is a 'relevant consideration' in determining whether a reasonable 'fit' exists." Michael W. Field, *On Tap,* 44 Liquormart, Inc. v. Rhode Island: *Last Call for the Commercial Speech Doctrine*, 2 ROGER WILLIAMS U. L. REV. 57, 70 (1996) (citations omitted).

134. Posadas de Puerto Rico Assocs. v. Tourism Co. of P.R., 478 U.S. 328 (1986).

135. *Id.*

136. *Id.* at 348.

137. *Id.* at 341. The Puerto Rican legislature believed that casino gambling among the local residents would increase crime, prostitution, and corruption. *Id.*

138. *Id.* at 341-44. With respect to whether the advertising restrictions directly advanced the government's interest, the Court stated "The Puerto Rico Legislature obviously believed . . . that advertising of casino gambling aimed at the residents of Puerto Rico would serve to increase the demand for the product advertised. We think the legislature's belief is a reasonable one. . . ." *Id.* at 341-42. The Court also deferred to the legislature's judgment that the regulations were no more extensive than necessary. The Court held that it was up to the legislature to decide if other measures would be as effective in furthering the state's interest as an advertising restriction. *Id.* at 344. *But see* Edenfield v. Fane, 507 U.S. 761, 770-71 (1993) (in striking down a ban on in-person solicitation by CPAs, the Court did not defer to the government when it held that the Board of Accountancy failed to sustain its burden of showing that the restriction directly advanced the asserted substantial interest).

139. In the *Posadas* decision, the Court abandoned its "least restrictive alternative analysis" in favor of a "rational basis" standard of review for the fourth prong. David D. Vestal, *The Tobacco Advertising Debate: A First Amendment Perspective*, ADVERTISING AND COMMERCIAL SPEECH 140, (Hon. Theodore R. Kupferman ed., 1990). "[T]he casino advertising ban . . . should have failed the fourth prong of *Central Hudson.*" *Id.* at 139. However, the Court adopted a "relaxed standard" because Puerto Rico was not required to show that other least restrictive alternatives would be ineffective. *Id.* at 140.

Perhaps even more damaging to the commercial speech doctrine is what Justice Rehnquist wrote after applying the *Central Hudson* analysis. In apparently adopting the "greater-includes-the-lesser" rationale,[140] Justice Rehnquist wrote that "it is precisely because the government could have enacted a wholesale prohibition of the underlying conduct that it is permissible for the government to take the less intrusive step of allowing the conduct, but reducing the demand through restrictions on advertising."[141] This seems to "add[] another dimension to the test of restrictions on commercial speech . . . if the underlying conduct is not protected, the [government] apparently has great latitude in the regulation of advertising about it."[142] In *United States v. Edge Broadcasting Co.*,[143] the Court held that "the activity underlying the relevant advertising—gambling—implicates no constitutionally protected right; rather, it falls into a category of 'vice' activity that could be . . . banned altogether."[144] Thus, as this theory applies to tobacco, unquestionably a vice product, if the government has the authority to completely ban the sale of tobacco products, then it implicitly has the authority to restrict its advertising and promotion. While the FDA asserts that it has the authority to completely ban the sale of tobacco,[145] neither Congress, nor any court, has given it that authority.[146]

2. 44 Liquormart, Inc. v. Rhode Island.[147]—A 1996 Supreme Court decision seems to have disavowed a majority of the *Posadas* opinion and afforded commercial speech its most protection since *Virginia Board of Pharmacy*. In *44 Liquormart*, all nine justices agreed that a state ban on advertising the prices of liquor was unconstitutional.[148] Although it was a plurality opinion, the

---

140. Under this rationale, the argument is that since the government has the greater power to ban the sale of a product, logically, that includes the lesser power to allow sales but with tighter restrictions. Martin H. Redish, *Tobacco Advertising and the First Amendment*, 81 IOWA L. REV. 589, 599-600 (1996) (arguing that the "greater-includes-the-lesser" rationale should not be a factor in First Amendment jurisprudence because its "'greater-includes-the-lesser' logic, when used in this context, actually stands the Constitution on its head.").

141. *Posadas*, 478 U.S. at 346.

142. Denise M. Trauth & John L. Huffman, *The Commercial Speech Doctrine*, ADVERTISING AND COMMERCIAL SPEECH 99, 111 (Hon. Theodore R. Kupferman ed., 1990).

143. 509 U.S. 418 (1993).

144. *Id.* at 426.

145. Harder, *supra* note 11, at 416 (citing 60 Fed. Reg. 41,355 (1995)). The FDA in its proposed rule stated that "it could have banned the sale or distribution of [tobacco products]." *Id.*

146. *Id.* at 415. Even if a court finds that the FDA has the authority to completely ban the sale of tobacco, *Posadas* is distinguishable because the FDA promulgated the regulations as a federal agency, where the regulation in *Posadas* was passed by the legislature. *Id.* at 416 n.117.

147. 116 S. Ct. 1495 (1996).

148. The Rhode Island ban prohibited vendors licensed in the state or out-of-state manufacturers, wholesalers, and shippers from "'advertising in any manner whatsoever' the price of any alcoholic beverage offered for sale in the State; the only exception is for price tags or signs displayed with the merchandise within licensed premises and not visible from the street." *Id.* at 1501. The ban also prohibited publication or broadcast of any advertisements that "make reference

importance of *44 Liquormart* rests in the laying of the foundation by the Justices to abandon sixteen years of precedent following *Central Hudson*.[149]

Delivering the principle opinion of the Court, Justice Stevens, joined by Justice Kennedy and Justice Ginsberg, rejected the argument that all commercial speech regulations are subject to the same level of review.[150] Justice Stevens stated that when advertising is regulated to protect consumers from deceptive and misleading advertising, then the regulation will be subject to "less than strict review."[151] However, when truthful, non-misleading messages are prohibited "for reasons unrelated to the preservation of a fair bargaining process, there is far less reason to depart from the rigorous review that the First Amendment generally demands."[152] Stevens proceeded to warn of the dangers created by governmental paternalism, stating that bans against truthful, non-misleading commercial speech usually occur because the government assumes "that the public will respond 'irrationally' to the truth."[153] However, "[t]he First Amendment directs us to be especially skeptical of regulations that seek to keep people in the dark for what the government perceives to be their own good."[154]

Justice Stevens, joined by Justices Kennedy, Souter, and Ginsberg, reviewing the ban under the "strict review" standard, held that despite Rhode Island's substantial interest in promoting temperance, the ban failed to directly advance the state's interest and it failed to satisfy the requirement that the restriction on speech be "no more extensive than necessary."[155] Justice Stevens stated that with respect to the fourth prong of the *Central Hudson* test, that if any non-speech alternative exists that would be more likely to achieve the state's substantial interest, then the regulation must fail.[156] Because non-speech alternatives almost always exist, the fourth prong of the *Central Hudson* test would rarely be satisfied.[157] Therefore, its usefulness may be short-lived.

Justices Stevens, Kennedy, Thomas, and Ginsberg also disavowed three critical aspects of *Posadas* which gave the government more freedom to regulate commercial speech. First, the Court held that *Posadas* "clearly erred in concluding that it was 'up to the legislature' to choose suppression over a less

---

to the price of any alcoholic beverages." *Id.*

149. Field, *supra* note 131, at 70.

150. *44 Liquormart*, 116 S. Ct. at 1507.

151. *Id.*

152. *Id.*

153. *Id.* at 1508.

154. *Id.*

155. *Id.* at 1510.

156. *Id.* To reduce consumption and promote temperance, Justice Stevens noted that higher prices (the effect the advertising ban would have had on alcohol) could be achieved by direct regulation or increased taxes, or that educational campaigns could be implemented to discourage consumption. *Id.* Applying strict review to the fourth prong will result in a "least restrictive means test," something that at least three members of the Court seem willing to do in commercial speech cases for the first time. Field, *supra* note 131, at 76.

157. Field, *supra* note 131, at 76.

speech-restrictive policy."[158] Second, the Court rejected the "greater-includes-the-lesser" theory because, contrary to the *Posadas* opinion, a ban on speech is sometimes more intrusive than banning conduct.[159] With respect to constitutional priorities, the Court noted that the right to free speech is valued more than conduct because of the "essential role that the free flow of information plays in a democratic society."[160] Finally, the Court rejected Rhode Island's argument that there should be a "vice" exception to the commercial speech doctrine.[161] The Court recognized the slippery slope of carving such an exception because of the difficulty in defining what constituted a vice activity. "Almost any product that poses some threat to public health or public morals might reasonably be characterized . . . as relating to a 'vice activity.'"[162] Furthermore, the effect of such an exception would allow the government to "justify censorship by the simple expedient of placing the 'vice' label on selected lawful activities . . . ."[163]

Justice Thomas, in his concurring opinion, was the most critical of the existing commercial speech doctrine. Thomas wrote that when the government attempts to keep legal users of a product ignorant in order to manipulate their choices in the marketplace, the *Central Hudson* test should not be applied; rather, the government's interest should be "per se illegitimate."[164] Thomas stressed the importance of free dissemination of information in a democratic society, the anti-paternalistic premises of the First Amendment, and the inappropriateness of manipulating consumer choices through the suppression of accurate commercial information.[165] Justice Thomas refused to join the principle opinion applying the *Central Hudson* test because he believed that it should not be used in this case.[166] He urged the Court to abandon *Central Hudson* and follow the doctrine in *Virginia Pharmacy Board* which states "that all attempts to dissuade legal choices by citizens by keeping them ignorant are impermissible."[167]

---

158. *44 Liquormart*, 116 S. Ct. at 1511.

159. *Id.* at 1512.

160. *Id.*

161. *Id.* at 1513.

162. *Id.*

163. *Id.; see also* Rubin v. Coors Brewing Co., 514 U.S. 476, 482 n.2 (1995) (rejecting the government's argument that legislatures have greater latitude in regulating speech "that promotes socially harmful activities, such as alcohol consumption, than they have to regulate other types of speech").

164. *44 Liquormart*, 116 S. Ct. at 1516 (Thomas, J., concurring).

165. *Id.* at 1517.

166. *Id.* at 1518. Thomas noted that both Justices Stevens and O'Connor adopted a stricter approach in applying the fourth prong of the *Central Hudson* test than in previous opinions. Thomas stated that these opinions "commit the courts to striking down restrictions on speech whenever a direct regulation (i.e., a regulation involving no restriction on speech regarding lawful activity at all) would be an equally effective method of dampening demand by legal users." *Id.* at 1519. Thomas concluded that "virtually all restrictions with such a purpose would fail the fourth prong of the *Central Hudson* test." *Id.*

167. *Id.* at 1520.

Justice O'Connor's concurring opinion, joined by Chief Justice Rehnquist, and Justices Souter and Breyer, decided this case "more narrowly" by applying the *Central Hudson* test.[168] Nevertheless, O'Connor held that the advertising ban failed the fourth prong because the fit between Rhode Island's method and its goal was not reasonable.[169] O'Connor's opinion implicitly adopts a stricter standard for the fourth prong of the *Central Hudson* test. O'Connor struck down the advertising ban because there were alternatives at the state's disposal which would have achieved the same goal without infringing on speech.[170] Thus, it seems that the Court will no longer hold that a method of impinging speech is a "reasonable fit" if there are alternatives available that do not intrude upon the ability to provide truthful, non-misleading information.

IV. THE BAN ON BRAND-NAME TOBACCO SPONSORSHIP IS UNCONSTITUTIONAL UNDER COMMERCIAL SPEECH JURISPRUDENCE

While there is little question that *44 Liquormart* has weakened the vitality of *Central Hudson* ,[171] the Court has not adopted a new analytical framework for commercial speech. Therefore, Part Four of this Note will use *Central Hudson* as the principle analytical tool in assessing the constitutionality of the FDA's restrictions on brand-name tobacco sponsorship, keeping in mind the impact *44 Liquormart* has had on the commercial speech doctrine.

A. Brand-Name Tobacco Sponsorships Are Lawful and Not Misleading

Brand-name tobacco sponsorships must be lawful and not misleading in order to receive First Amendment protection.[172] Tobacco is a legal product in the United States, and the sponsorship of events is a legal activity. On its face, event sponsorship by brand-name tobacco products appears not to be deceptive or misleading. For example, in the NASCAR Winston Cup Series, Winston does not make an attempt to promote cigarettes by making any claims about the benefits of smoking, or stating reasons why a person should buy Winston cigarettes. The Winston name and logo is what is being promoted. Furthermore, the advertisement of tobacco is no more misleading or deceptive than the advertisement of many other products. For example, tobacco advertisements do not mention the adverse consequences associated with tobacco use; however, butter manufacturers do not mention the adverse health consequence of butter in their advertisements either.[173] Thus, the commercial message that is put forth is not deceptive or misleading, and should receive First Amendment protection.

The FDA argues that since it is illegal in all fifty states for persons under the

---

168. *Id.* at 1521.

169. *44 Liquormart*, 116 S. Ct. at 1521 (O'Connor, J., concurring).

170. *Id.*

171. *See generally* Field, *supra* note 131 (arguing that the Court is ready to abandon *Central Hudson* and that commercial speech should be afforded full First Amendment protection).

172. *Central Hudson*, 447 U.S. at 563-64.

173. Stoner, *supra* note 44, at 654.

age of eighteen to purchase tobacco, then the advertisement of tobacco products is not a lawful activity when advertisements are directed towards persons under the age of eighteen.[174] In making its argument, the FDA relies on the well-established principle that commercial speech "related to" unlawful activity does not merit First Amendment protection.[175] The FDA asserts that tobacco advertising is "related to" an illegal activity in two respects.[176] First, tobacco advertisements propose a commercial transaction that do not differentiate between adult and minor purchasers.[177] Therefore, because it is unlawful for minors to purchase tobacco products in all fifty states, the undifferentiated tobacco advertisements, "at least in part," are unlawful.[178] Second, tobacco advertisements are "related to" an unlawful activity. Advertising has a "powerful appeal" to children and affects their decision to use tobacco, which generates an attempt to purchase cigarettes.[179] Because it is unlawful for minors to purchase cigarettes, tobacco advertising "can appropriately be viewed as encouraging, and thus being 'related to' an illegal activity."[180] The FDA contends that it should be afforded the discretion to differentiate between advertising that "relates to" children, which it claims is unlawful, and advertising that does not.[181]

The FDA's reasoning does not stand on solid ground. With respect to the argument that tobacco advertisements are unlawful, the FDA provided no solid evidence for support of this claim. In fact, the only data the FDA provided to support this contention was the number of cigarettes that children and adolescents smoke each year.[182] Because of these statistics, the FDA asserts that

---

174. FDA Regulations, *supra* note 3, at 44,471.

175. *Id.* (citing *44 Liquormart*, 116 S. Ct. at 1505 n.7) ("By contrast, the First Amendment does not protect commercial speech about unlawful activities."); Florida Bar v. Went For It, 515 U.S. 618, 623-24 (1995) ("Under *Central Hudson*, the government may freely regulate commercial speech that concerns unlawful activity or is misleading."); Bolger v. Youngs Drug Prods. Corp., 463 U.S. 60, 69 (1983) ("The State may also prohibit commercial speech related to illegal behavior."); *Central Hudson*, 447 U.S. at 563-64 ("The government may ban . . . commercial speech related to illegal activity." (citations omitted)).

176. FDA Regulations, *supra* note 3, at 44,471.

177. *Id.*

178. *Id.*

179. *Id.*

180. *Id.*

181. *Id.* The FDA relied on Justice Stevens' opinion in *44 Liquormart* which distinguished *United States v. Edge Broadcasting Co.*, 509 U.S. 418 (1993). The Court in *Edge* upheld a statute which prohibited lottery advertisements by a broadcaster who was licensed in a state where lotteries were illegal. *Edge Broad.*, 509 U.S. at 435. In *44 Liquormart*, Justice Stevens stated that the statute in *Edge* "was designed to regulate advertising about an activity that had been deemed illegal in the jurisdiction in which the broadcaster was located," as compared to *44 Liquormart*, where the statute "targets information about entirely lawful behavior." FDA Regulations, *supra* note 3, at 44,472 (citing *44 Liquormart*, 116 S. Ct. at 1511). The FDA contends that this is the same type of distinction that it is drawing with respect to tobacco advertising. *Id.*

182. *Id.* at 44,471.

in a "practical sense" tobacco advertising is unlawful.[183] However, the fact that children are consuming an extraordinary amount of tobacco each year does not make tobacco advertising unlawful. By providing statistics on the number of underage smokers, the FDA underscores a serious problem in society; it does not, however, provide a nexus between the prevalence of underage smoking and the illegality of tobacco advertisements.

The FDA also faults the tobacco industry for not differentiating between adults and minors in advertisements. The failure to do so, according to the FDA, results in an unlawful activity when the product is advertised. This reasoning is also faulty. First, it is the *sale* of tobacco to minors that is illegal, not the advertisement of tobacco products. Second, it is absurd to suggest that tobacco manufacturers are obligated to expressly differentiate between adults and children in its advertisements. Is the FDA suggesting that tobacco manufacturers must explicitly state in its advertisements that the message is only targeted to the lawful users of the product? Beer advertisements do not. Does this mean that every beer commercial on television is illegal? Under the FDA's logic, it seems that they would be. Beer commercials or lottery advertisements would be unlawful activities because only adults can drink alcohol or play the lottery, yet those advertisements do not attempt to differentiate between minors and adults.

The greater effect of the FDA's logic would be to make the advertisement of nearly every "vice" product an unlawful activity. Most vice activities are restricted to adults, and under the FDA's reasoning, advertising of such products would be unlawful and therefore could be heavily regulated. This is in direct contradiction to the Court's ruling in *44 Liquormart* that "vice activities" are not afforded less protection[184] and with the well-established position that the Court has taken against governmental paternalism.[185]

Although advertisements that promote adult activities may be appealing to

---

183. *Id.*

184. *See supra* Part III.C; *see also* Mark R. Ludwikowski, *Proposed Government Regulation of Tobacco Advertising Uses Teens to Disguise First Amendment Violations*, 4 COMMLAW CONSPECTUS 105, 110 (1996) ("There is little merit in asserting that First Amendment protection should not be accorded commercial speech that advertises a legal but harmful product.").

185. *44 Liquormart*, 116 S. Ct. at 1517 n.2 (Thomas, J., concurring); *see also* Linmark Assocs., Inc. v. Township of Willingboro, 431 U.S. 85, 96-97 (1977); Bates v. State Bar of Ariz., 433 U.S. 350, 364-65 (1977); Friedman v. Rogers, 440 U.S. 1, 8-9 (1979); Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n, 447 U.S. 557, 561-62 (1980); Bolger v. Youngs Drug Prods. Corp., 463 U.S. 60, 79 (1983) (Rehnquist, J., for two Justices, concurring); Zauderer v. Office of Disciplinary Counsel, 471 U.S. 626, 646 (1985); Posadas de Puerto Rico Assocs. v. Tourism Co. of P.R., 478 U.S. 328, 350-51 (1986) (Brennan, J., for three Justices, dissenting); Cincinnati v. Discovery Network, Inc., 507 U.S. 410, 421-22 n.17 (1993); Edenfield v. Fane, 507 U.S. 761, 767, 770, (1993); United States v. Edge Broad. Co., 509 U.S. 418, 437-39 & nn.1, 3 & 4 (1993) (Stevens, J., for two other Justices, dissenting); Ibanez v. Florida Dept. of Business & Professional Regulation, 512 U.S. 136, 140-45 (1994); Rubin v. Coors Brewing Co., 514 U.S. 476, 481 (1995) (Stevens, J., concurring); Florida Bar v. Went For It, Inc., 515 U.S. 618, 638-39 (1995) (Kennedy, J., for four Justices, dissenting).

children and adolescents, as well as to adults, this is not a sufficient reason to justify a sponsorship ban. In *Bolger v. Youngs Drug Products Corporation*, the Court held that "the government may not 'reduce the adult population . . . to reading only what is fit for children.'"[186] The tobacco industry noted that it is "the cartoon form of Joe Camel that causes people to mistakenly believe that Joe Camel is child-oriented."[187] Many adult-oriented products use cartoon figures to promote their products, like "the Pink Panther for fiberglass insulation, Garfield the Cat for a hotel chain, Mr. Clean for household products, and the Peanuts characters for life insurance."[188] The Joe Camel campaign has proven effective in reaching the eighteen to twenty-four year-old audience, an audience which can lawfully purchase tobacco.[189] Prohibiting advertisements of potentially harmful products on the basis that they are appealing to children will produce inconsistent and unpredictable results. "It is not improbable to suspect that a ban on tobacco advertising will lead to gags on manufacturers of other products that at any given time may be considered politically incorrect."[190]

The FDA should not be afforded the discretion to determine which advertising "relates to" children. The danger of giving the FDA discretion to make such a determination is apparent from its regulations on brand-name sponsorship. Those regulations fail to make any distinction with respect to whether the event consists of primarily an adult audience or an audience with a substantial number of children. The brand-name sponsorship ban applies equally to all events. Thus, either the FDA is saying all sponsored events "relate to" children and adolescents, which is clearly wrong,[191] or the FDA has shown that it is unable to determine which advertising "relates to" children and adolescents so they imposed a complete brand-name sponsorship ban. Either way, the FDA has demonstrated that it should not be afforded the discretion to make this determination.

Although the FDA did not expressly state that tobacco advertising is misleading, they did hint that it might be.[192] The FDA stated that children and adolescents are "very impressionable and therefore vulnerable to the sophisticated marketing techniques employed by the tobacco industry, techniques that associate the use of tobacco products with excitement, glamour, and independence."[193] Nevertheless, it is unlikely that a court will find it to be so.[194]

---

186. Bolger v. Youngs Drug Prods. Corp., 463 U.S. 60, 73 (1983) (quoting Butler v. Michigan, 352 U.S. 380, 383 (1957)).

187. FDA Regulations, *supra* note 3, at 44479.

188. *Id.*

189. *Id.* Joe Camel advertisements were placed in *Cycle World*, *Penthouse*, *Gentleman's Quarterly*, and *Road and Track*. *Id.* Camel's market share among 18 to 24 year-olds increased from 3.2% in 1986, the year before the inception of Joe Camel, to 10.1% in 1994. *Id.*

190. Ludwikowski, *supra* note 178, at 110.

191. *See* discussion *infra* Part IV.D.

192. FDA Regulations, *supra* note 3, at 44,398.

193. *Id.*

194. Ludwikowski, *supra* note 178, at 111.

It would be difficult to identify the criteria to determine whether advertisements were aimed at seventeen year-olds, who are prohibited from purchasing tobacco, or at nineteen year-olds, who can legally purchase tobacco.[195] Furthermore, the requirement of warning labels on packages effectively counters "any misleading effects of tobacco advertising."[196] The FDA did not predicate their regulation of tobacco advertising solely on it being unlawful or misleading, therefore the regulations must satisfy the remaining three prongs of the *Central Hudson* test.

B. The FDA's Interest Is Substantial

The FDA's interest in protecting children and adolescents from the hazards of tobacco is substantial.[197] The cost of smoking-related illnesses and death were calculated to be in excess of $68 billion in 1990, including $20.8 billion in direct health care costs, $6.9 billion in morbidity costs, and $40.3 billion in lost future earnings due to premature death.[198] Furthermore, it is estimated that one million persons under the age of eighteen start smoking each year.[199]

The FDA's asserted interest in protecting children and adolescents is consistent with other interest which courts have held to be substantial. These include energy conservation,[200] esthetics of a city,[201] and the ill effects of gambling on residents.[202] Considering the serious consequences of smoking and the great costs it imposes on society, the aforementioned interests pale in comparison to the substantial interest in reducing underage smoking. Additionally, in *Penn Advertising of Baltimore, Inc. v. Mayor and City Council of Baltimore*,[203] the Fourth Circuit held that "reducing cigarette consumption by

---

195. *Id.*

196. *Id.*

197. The title of the FDA's rule is *Regulations Restricting the Sale and Distribution of Cigarettes and Smokeless Tobacco to Protect Children and Adolescents*. However, in the text of the rule, the FDA states the substantial interest as "protecting the public health." FDA Regulations, *supra* note 3, at 44,472. While this may seem to be an insignificant distinction because the health of the general population would seem to be as much of a substantial interest than that of children, this statement seems to reveal that the FDA is less concerned with underage use of tobacco products than it is with tobacco use in general. However, by asserting that the protection of children and adolescents as their substantial interests, the FDA can effectively avoid the problem of being too paternalistic. *See, e.g.*, Action for Children's Television v. FCC, 58 F.3d 654, 661 (D.C. Cir. 1995); New York v. Ferber, 458 U.S. 747, 756-57 (1982); Denver Area Educ. Telecommunications Consortium, Inc. v. FCC, 64 U.S.L.W. 4706 (1996).

198. FDA Regulations, *supra* note 3, at 44572 (citing Statement of Clyde Behney and Maria Hewitt on Smoking-Related Deaths and Financial Costs: Office of Technology Assessment Estimates for 1990 Before the Senate Finance Committee 2 (April 28, 1994)).

199. *Id.*

200. *Central Hudson*, 447 U.S. at 569.

201. Cincinnati v. Discovery Network, 507 U.S. at 416.

202. *Posadas*, 478 U.S. at 341.

203. 63 F.3d 1318 (4th Cir. 1995), *vacated sub nom.* Penn Advertising, Inc. v. Schmoke, 116

minors constitutes a substantial public interest."[204] Furthermore, that the sale of tobacco products is illegal in all fifty states strongly suggests that the interest in underage smoking is more than substantial, it is irrefutable.[205] Therefore, the second prong of the *Central Hudson* test is satisfied.

C. The FDA's Regulations May Not Directly Advance Its Asserted Interest

The FDA's contention that the advertising restrictions will directly advance their interest stands on tenuous grounds. Under the third prong of the *Central Hudson* test, the FDA must prove that the advertising restrictions will directly advance their asserted interest to a "material degree."[206] While the Court has shown great deference to the legislature in years past,[207] the Court has departed from that approach and now seems to require evidence which is more than "speculation or conjecture,"[208] especially when the government "takes aim at accurate commercial information for paternalistic ends."[209] The FDA's regulation seems to be well-supported by evidence that advertisements directly impact the number of new underage smokers; however, a closer look at the FDA's empirical data reveals that their conclusion is based primarily on speculation or conjecture.

The FDA acknowledged that no one study or piece of evidence would prove that the advertising restrictions would significantly decrease tobacco use by minors.[210] Instead, the FDA broke down the effects of advertising on children and adolescents into several components.[211] According to the FDA, when considered together, the evidence proved that the restrictions on advertising directly advanced the government's interest.[212] In reality, the FDA provided

---

S. Ct. 2575 (1996); *adopted as modified*, Penn Advertising, Inc. v. Mayor & City Council, 101 F.3d 332 (4th Cir. 1996).

204. *Id.* at 1325.

205. Harder, *supra* note 11, at 418.

206. *44 Liquormart*, 116 S. Ct. at 1509. For purposes of analyzing this prong, a reviewing court will probably consider the cumulative effect of all of the FDA's advertising restrictions. Unlike prior cases where there was essentially one advertising restriction, e.g., a prohibition against the advertising of alcohol prices, the FDA's rule imposes several advertising restrictions in the various media. A court could require that each individual restriction, e.g., the prohibition against brand-name tobacco sponsorships, directly advances the government's interest. However, the more plausible solution would be to analyze the third prong considering the effect the entire regulation will have, and then analyze each individual provision separately under the fourth prong of the *Central* Hudson test. This Note utilizes this approach.

207. *See supra* Part III.C.

208. *44 Liquormart*, 116 S. Ct. at 1510.

209. *Id.*

210. FDA Regulations, *supra* note 3, at 44,476.

211. *Id.* at 44,475.

212. *Id.*

piece-mealed evidence that, when analyzed closely, proves very little.[213]

First, the FDA stated that "perhaps the most compelling evidence" that advertising affects a young person's decision to use tobacco is that tobacco is among the most heavily advertised products in America.[214] The FDA failed to show a nexus between the size of the industry's advertising budget and an increase in tobacco use by children. Nevertheless, the FDA claimed that this evidence demonstrates that advertising creates a "friendly familiarity" that makes smoking seem "respectable to young people."[215] This conclusion is too general. Other factors play an integral part in whether a child perceives smoking as respectable and, ultimately, whether a child decides to smoke. A child whose parents and peers express negative views about smoking is less likely to view smoking as "respectable," despite the prevalence of tobacco advertisements.

Second, the FDA cited studies which showed "that children who smoke are more likely to correctly identify cigarette advertisements and slogans in which the product names or parts of the slogans have been removed than are children who do not smoke."[216] While the FDA acknowledged that these studies did not establish that exposure to, recall of, approval of, and response to advertising caused children to smoke, the FDA included these studies because they showed that advertisements created "an important role in developing an appealing and memorable image for brands."[217] However, "an appealing and memorable image for brands" does not necessarily equate to an increase in the number of underage smokers. The subjects of the study were children who already smoked. While this study may indicate that existing smokers may switch brands of tobacco because of advertising exposure, it proves nothing with respect to the effects of advertising on children who do not smoke.

Third, studies showed that cigarette advertising caused children to overestimate the prevalence of smoking in society.[218] While the studies did not show a causal relationship between the overestimation of smoking and the number of children who start using tobacco because of the overestimation, the FDA stated that the studies were included to show the "acceptability" of smoking.[219] Once again, this item of evidence does nothing to establish that the advertising restrictions would advance the government's interest.

Fourth, evidence was presented to show the effectiveness of ad campaigns with respect to children.[220] One study showed that ninety-one percent of six year-olds and thirty percent of three year-olds recognized "Joe Camel," the cartoon

---

213. The following analysis discusses the majority of the FDA's reasons why the regulations satisfy the third prong of *Central Hudson* and is a fairly complete portrayal of their assertions. However, some of the less persuasive rationales were omitted.

214. *Id.* at 44,475.

215. *Id.*

216. *Id.*

217. *Id.* at 44,476.

218. *Id.*

219. *Id.*

220. *Id.* at 44,477 n.110.

character marketing Camel cigarettes.[221] As with the other categories of evidence, this information was included not to show any correlation between advertising and smoking, but rather to show the "pervasiveness of tobacco advertising."[222]

Fifth, the FDA presented evidence of an internal memo from the tobacco industry specifically addressing the issue of targeting young people.[223] While this information is certainly damaging to the tobacco industry, and the "logical inference" is that advertising "play[s] an important role in young people's smoking behavior,"[224] this evidence does not show that the advertising restrictions would directly advance the government's interest "to a material degree."

Sixth, the FDA offered studies to show that advertising affects the brand choices of underage tobacco users.[225] This item of evidence also seems irrelevant to the establishment of the FDA's proposition because any affect on brand preference would only redistribute the market share between the tobacco companies. The FDA did not show any correlation between brand choice and an increase in the number of children who smoke. When considered together, all the FDA has established is that tobacco advertising is pervasive in society and that children are aware of it. Proving that children are alert to advertisements and are able to recognize slogans and cartoon characters does not establish that advertising restrictions will decrease tobacco consumption by minors.

Nevertheless, the FDA may still be able to satisfy its burden in showing that the regulations directly advance its asserted interest. A 1994 report issued by the Surgeon General concluded that "[a] substantial and growing body of scientific literature has reported on young people's awareness of, and attitudes about, cigarette advertising . . . [and when] . . . considered together, these studies offer a compelling argument for the mediated relationship of cigarette advertising and adolescent smoking."[226] However, two major problems exist for relying on this report to satisfy the third prong of *Central Hudson*. The first is that the Surgeon General found only a "compelling argument" for the nexus between advertising and consumption.[227] In fact, in 1989, the Surgeon General stated that there was

---

221. *Id.* (citing P.M. Fischer et al., *Brand Logo Recognition by Children Aged 3 to 6 Years: Mickey Mouse and Old Joe the Camel*, 266 JAMA 3145, 3145-48 (1991)).

222. *Id.*

223. *Id.* at 44,480. The memo reads, in relevant part:

> Evidence now available . . . indicate[s] that the 14 to 18 year old group is an increasing segment of the smoking population. RJR must soon establish a successful new brand in this market if our position in the industry is to be maintained over the long-term."

Id. at 44,481.

224. *Id.* at 44,482.

225. *Id.*

226. U.S. Dep't of Health and Human Servs., *Preventing Tobacco Use Among Young People: A Report of the Surgeon General*, 188 (1994).

227. *Id.*

no study available which would provide a definitive answer to whether there is a link between advertising and youth consumption, nor would it be likely that one would be "forthcoming in the foreseeable future."[228] Secondly, assuming *arguendo*, that the Surgeon General is correct that advertising does play a role in youth tobacco use, because the extent is unknown, it is impossible to prove that an advertising ban will reduce underage tobacco use to a "material degree."

Perhaps the most compelling evidence which the FDA presented was empirical data from studies conducted in other countries where tobacco advertising had been banned. Studies showed that after an advertising restriction was put in place, the percentage of teenagers who smoked decreased significantly.[229] The FDA also noted that the Court has recognized the relationship between advertising and demand for a product in recent decisions.[230] However, in *44 Liquormart*, the Court refused to hold that Rhode Island's interest was advanced when the evidence showed the advertising restriction would produce only a "marginal impact" on consumption.[231] The Court held that the result of the regulation must be "significant."[232] The FDA will be required to show that a ban on tobacco advertising will produce a significant reduction in tobacco use by minors in the United States as occurred in other countries. It is questionable whether such a showing will be made in the United States and whether the third prong will be satisfied.

D. The Ban on Brand-Name Tobacco Sponsorships is Not Narrowly Drawn

The FDA's ban on brand-name tobacco sponsorships cannot survive constitutional review because it fails the fourth prong of the *Central Hudson* test. The Court's most recent commercial speech decision, *44 Liquormart*,[233] presents a substantial obstacle which the regulations cannot overcome. Despite the FDA's attempt to minimize that decision, *44 Liquormart* has reshaped the future of

---

228. Ludwikowski, *supra* note 178, at 113 n.115 (citing U.S. Dep't of Health and Human Servs., *Reducing the Health Consequences of Smoking: 25 Years of Progress; A Report of the Surgeon General*, 516-17 (1989).

229. In Norway, the percentage of fifteen-year-old boys and girls who smoked in 1975, before a restriction on tobacco advertising and promotion was put in place, was approximately 23% and 28%, respectively. FDA Regulations, *supra* note 3, at 44,490-91. In a 1986-87 follow-up survey, that percentage had decreased to 16% and 17% respectively. *Id.* Norway banned all advertising of tobacco products in 1975. *Id.* Between 1975 and 1990, the percentage of daily smokers aged 13 to 15 declined from 15% to 9% for boys and from 17% to less than 10% for girls. *Id.*

230. *Coors Brewing Co.*, 514 U.S. at 487 ("It is assuredly a matter of 'common sense' that a restriction on the advertising of a product characteristic will decrease the extent to which consumers select a product on the basis of that trait."); *44 Liquormart*, 116 S. Ct. at 1506 (Justice Stevens quoted with apparent approval of *Central Hudson's* reliance on the "immediate connection" between "promotional advertising" and demand).

231. *44 Liquormart*, 116 S. Ct. at 1510.

232. *Id.* at 1509-10.

233. *See supra* Part III.C.

commercial speech jurisprudence, and cannot be ignored.

The FDA attempts to circumvent *44 Liquormart* by (1) relying on precedent which has been either disavowed or subsequently modified; (2) distinguishing *44 Liquormart*; and (3) asserting that the regulations are still consistent with commercial speech jurisprudence.

First, the FDA stated that the Court does not use the "least restrictive means" test; rather all that is necessary is a "reasonable fit" between the regulation and the government's substantial interest to satisfy the fourth prong of *Central Hudson*.[234] However, the FDA failed to consider that in *44 Liquormart* the Stevens bloc seemed to implicitly adopt the "least restrictive means" test;[235] the O'Connor bloc found that less burdensome alternatives may indicate that the fit is not reasonable;[236] and that Justice Thomas wrote that the *Central Hudson* test should be abandoned completely.[237] Clearly, the FDA's position that it had not "mischaracterized its burden"[238] is wrong.

Second, the FDA asserted that the amount of constitutional protection afforded commercial speech is "commensurate with its subordinate position in the scale of First Amendment values," despite language to the contrary in *44 Liquormart*.[239] In *44 Liquormart*, three Justices stated that "when a State entirely prohibits the dissemination of truthful, nonmisleading commercial messages for reasons unrelated to the preservation of a fair bargaining process, there is far less reason to depart from the rigorous review that the First Amendment generally demands."[240] The FDA said that this statement had no application to the tobacco regulations because the "FDA is not entirely prohibiting the dissemination of commercial messages about cigarettes and smokeless tobacco . . . [and because] . . . the restrictions are related to the bargaining process."[241]

The FDA is incorrect on both accounts. The ban on brand-name sponsorship is effectively a complete ban. In *44 Liquormart*, Justice Stevens wrote that "Rhode Island's price advertising ban [on alcoholic beverages] constitute[d] a blanket prohibition against truthful, nonmisleading speech about a lawful product."[242] Rhode Island's advertising ban only prohibited the advertising of prices of alcoholic beverages, but allowed alcoholic beverages to be advertised in general.[243] Yet, the Court held that the advertising restriction was a "blanket

---

234. FDA Regulations, *supra* note 3, at 44,496. (citing State Univ. v. Fox, 492 U.S. 469, 480 (1989)).

235. *See supra* note 152 and accompanying text.

236. *See supra* note 170 and accompanying text.

237. *See supra* note 167 and accompanying text.

238. FDA Regulations, *supra* note 3, at 44,496.

239. *Id.* at 44,470. (citing Florida Bar v. Went For It, Inc., 515 U.S. 618, 623 (1995)).

240. *44 Liquormart*, 116 S. Ct. at 1507.

241. FDA Regulations, *supra* note 3, at 44,470.

242. *44 Liquormart*, 116 S. Ct. at 1508.

243. *Id.* at 1501.

prohibition."[244] The FDA's ban on brand-name sponsorship is much more intrusive. By allowing sponsorship only in the corporate name, a tobacco manufacturer cannot provide information to consumers about a particular product. Thus, the industry is completely banned from disseminating a commercial message about its product to consumers in the sponsorship context. The FDA's assertion that tobacco manufacturers can still provide information about a product is incorrect. If the product is laundry detergent, and the same rule is in effect, i.e. sponsorship only in the corporate name, Proctor & Gamble, a company that manufacturers many different brands of laundry detergent, could not effectively advertise the brand Tide if it was allowed only to sponsor an event under its corporate name. Therefore, the FDA's contention that its regulations are not a complete ban on advertising is without merit.

Furthermore, the FDA seems to mischaracterize what the Court in *44 Liquormart* meant when it referred to the "preservation of a fair bargaining process."[245] The Court sought to protect consumers from misleading, deceptive, or aggressive sales practices which would prevent them from making a meaningful choice.[246] The prohibition of truthful, nonmisleading advertising usually has the opposite effect by "hinder[ing] consumer choice."[247] Therefore, the stricter standard of review adopted in *44 Liquormart* cannot be dismissed as inapplicable.

Third, the FDA claims that no alternative to its regulation will directly advance the government's interest; therefore, the regulation is narrowly drawn.[248] The FDA contends that no governmental program has successfully reduced teenage smoking in the past; that non-speech alternatives are being implemented in conjunction with the speech restrictions; and for the government's interest in protecting children and adolescents to be furthered, restrictions on advertising must be imposed concurrently with non-speech restrictions.[249] Nevertheless, the ban on brand-name sponsorships is "more extensive than necessary." As the principle opinion in *44 Liquormart* suggests, the existence of alternative non-speech regulations will make the regulation fail the fourth prong.

Alternative forms of non-speech regulation are in abundance. Better enforcement of existing laws prohibiting the sale of tobacco to minors would directly advance the FDA's interest by reducing underage tobacco use. In 1992, Congress passed the Alcohol, Drug, Abuse, and Mental Health Administration (ADAMHA) Reorganization Act of 1992.[250] The Act prohibits the Department of Health and Human Services (DHHS) from providing block grants for the treatment and prevention of substance abuse to a state unless the state prohibits

---

244. *Id.* at 1508.

245. *Id.* at 1507.

246. *Id.*

247. *Id.* at 1508.

248. FDA Regulations, *supra* note 3, at 44,499.

249. *Id.*

250. ADAMHA Reorganization Act of 1992, Pub.L. No. 102-321, 106 Stat. 323 (1992). 42 U.S.C. § 201 (1994).

the sale and distribution of tobacco products to persons under eighteen.[251] The DHHS said that "[e]liminating virtually all sales [of tobacco products] to minors does not even present particularly difficult enforcement problems."[252]

An educational campaign warning children and adolescents about the associated risks of tobacco use would be another "less burdensome alternative."[253] An increased tax imposed upon tobacco would also likely reduce tobacco use. In *44 Liquormart*, Justice Stevens approved of increased taxation and educational campaigns as alternatives to restricting speech in striking down the alcohol price advertising ban. [254]

Providing warnings about the dangers associated with tobacco use at sponsored events would be less restrictive. For example, anywhere the name of a tobacco product appears at a race track (i.e. cars, flags or race programs) an appropriate warning label, similar to what is currently on tobacco products, could be displayed. Considering the available alternatives and the Court's stricter application of the fourth prong,[255] it seems that the regulations fail to pass constitutional muster on these factors alone.

Even more fatal to the sponsorship ban is that it does not differentiate between events attended primarily by adults and those attended by a substantial number of children. The ban is a blanket ban on all events. If a substantial number of children are not present or are not watching a sponsored event, like an automobile race, then the government's interest is not substantially advanced and the regulations are not narrowly drawn.

NASCAR is an adult-oriented sport. According to NASCAR statistics, only three percent of the spectators at NASCAR races are under the age of eighteen.[256] With an average attendance of 171,830, the regulations will "further" the FDA's substantial interest by protecting only 5,155 children and adolescents from tobacco advertising at each race. The FDA dismissed the fact that the number of children who attend races was not substantial. The FDA said that they "did not receive any data to support or refute these numbers," and that, in any event, "recent reports in the press indicate that the number of young people attending these events *may* be growing."[257] Instead, the FDA emphasized the glamour and

---

251. 42 U.S.C. §300x-26(a)(1) (1994).

252. 58 Fed. Reg. 45156, 45165 (1993).

253. As part of the rules, the FDA is not requiring an educational campaign. However, the FDA does plan to implement a educational campaign using the notification system under section 518(a) of the Federal Food, Drug, and Cosmetic Act (FDCA). If the FDA finds that a device "presents an unreasonable risk of substantial harm, then, after consultation with tobacco manufacturers, they can issue a notification that requires tobacco manufacturers to notify young people about the substantial health risks associated with tobacco. FDA Regulations, *supra* note 3, at 44,538.

254. *44 Liquormart*, 116 S. Ct. at 1510.

255. *See supra* Part III.C.

256. NASCAR PUBLIC RELATIONS MATERIALS, *supra* note 29.

257. FDA Regulations, *supra* note 3, at 44528 (emphasis added). Apparently the FDA did not make an attempt to verify the data; rather, they deferred to what was reported in the press.

excitement that is associated with sponsors of racing events or teams, even though that has no relevance to the regulations being narrowly drawn.[258]

The FDA also observed that, besides the spectators at a race, there are the millions of viewers watching on television.[259] While this is a valid consideration, the number of children who watch racing on television is not substantial. Children and adolescents only comprise seven percent of the total television audience. Combining spectators and television viewers, children and adolescents comprise only 6.7 percent of the total audience.[260]

For the FDA to claim that this number justifies a sponsorship ban is contrary to its own methodology in its restrictions on the use of color and imagery in print publications. To draw its print advertising restrictions as narrowly as possible, the FDA decided not to limit advertisements to a text-only format where the publication was primarily an "adult publication."[261] The FDA defines adult publications as those publications "(1) [w]hose readers age 18 or older constitute 85 percent or more of the publication's total readership, or (2) that are read by fewer than 2 million people under the age of 18, whichever method ensures the fewest young readers."[262] The FDA stated that their "concern is with advertising that affects minors and with tailoring the restrictions in this final rule to burden as little speech as possible" and "that an exception from the text-only requirement for publications that are read primarily by adults is still reasonable and feasible."[263]

With respect to imposing a similar type of threshold test to sponsored events, the FDA rejected the idea and stated that it was "not aware of any way to limit the restriction to events that are attended by young people."[264] Perhaps the FDA realized that under this type of analysis, motorsport events, as well as most other sponsored events, would fall outside the regulation. In fact, the number of adult spectators at a NASCAR Winston Cup race barely exceed the print regulations two million person benchmark.[265] Children and adolescents do not even come

---

258. *Id.* at 44,529.

259. *Id.* at 44,528.

260. Combining 5155 children who are live spectators and 180,791 children who are watching on television (calculated by dividing 64 million children (the total of children television viewers per year) by 354 (the number of televised events)), the average number of children and adolescents in affected at any one race is 185,946. The average total audience for any given race is 2,756,575 persons (2,584,745 television viewers and 171,830 spectators).

261. FDA Regulations, *supra* note 3, at 44,513.

262. *Id.*

263. *Id.* at 44,514.

264. *Id.* at 44,534.

265. The number of adults at a single NASCAR Winston Cup race is 166,675 (97% of 171,830 total spectators). The number of adult television viewers is 2,403,954 per race (93% of 915 million total television viewers, divided by 354 televised racing events). The total adult audience for a single race equals 2,570,629. The black and white, text-only format for print advertisements only takes effect when the number of readers under age 18 for the publication reaches two million. Therefore, the total number of adults who view in person and on television

close to reaching these numbers. Furthermore, under the two million readership benchmark, one weekly publication with an average youth readership of 1.5 million per week will yield a higher number of children being exposed to colorful and image-filled advertisements in one year than in one year of motorsport broadcasts. This is just one of many publications; the cumulative effect is far greater.

Using the FDA's own statistics and methodology, there is no question that NASCAR, and the entire world of motorsports, is primarily an adult industry. The FDA's regulations have effectively reduced what adults may see at a racetrack to "only what is fit for children."[266] Since the brand-name sponsorship ban serves to effectively prohibit the dissemination of truthful, nonmisleading advertisements to adults, the regulations must fail because they are not narrowly drawn.

V. PROPOSAL

While Parts III and IV of this Note observed that the Court in *44 Liquormart* seemed to imply that the existence of any non-speech alternative would cause the regulation to fail, the impact of that case is yet to be felt. The Court did not overrule *Central Hudson*; therefore, commercial speech cases are still decided by using a balancing test. Furthermore, the interpretation of *44 Liquormart* in the various circuits is still unknown. The Fourth Circuit, on remand from the Supreme Court to reconsider a case in light of *44 Liquormart*, upheld a city ordinance which prohibited the advertisement of tobacco on billboards and signs in a "publicly visible location" in designated locations.[267] If the fourth prong of the *Central* Hudson test is interpreted as strictly as *44 Liquormart* suggests, that the existence of non-speech alternatives make any speech regulation fail, then it is likely that the following proposal would fail as well. However, if the fourth prong does not require a true "least restrictive means," then the proposal will pass constitutional muster because it is narrowly drawn.

To ensure that it will pass constitutional muster, the ban on brand-name sponsorship of sporting events, teams, and entries should be redrafted similarly to the print regulations.[268] Brand-name sponsorship should be permitted when a substantial majority of the audience is adult. The FDA's fifteen percent/two

---

a single NASCAR race only exceeds the FDA's threshold by 570,629 persons. The number of children at a NASCAR race falls approximately 1.8 million persons under the threshold. *NASCAR Demographics*, *supra* note 29.

266. *Bolger*, 463 U.S. at 73 (striking down a ban which prohibited unsolicited advertisements for contraceptives).

267. Penn Advertising, Inc. v. Mayor & City Council, 101 F.3d 332 (4th Cir. 1996). In its analysis of the fourth prong, the court only looked at whether there were less restrictive means of screening outdoor advertising from minors, rather than if there were any less restrictive means which would directly advance the city's interest in protecting minors. The court found that there were no less restrictive means, and thus, the advertising satisfied the fourth prong. *Id.*

268. *See supra* Part IV.D.

million person benchmark for print advertisements is a reasonable method for determining whether brand-name sponsorships should be allowed. This number ensures that the advertising ban would affect a substantial number of children, and, thus, would further the government's interest in protecting children.

The new rule should read as follows:

> No manufacturer, distributor, or retailer may sponsor or cause to be sponsored any athletic, musical, artistic, or other social or cultural event, *series of events*, or any entry or team in any event *or series of events*, in the brand name, logo, symbol, motto, selling message, recognizable color, pattern of colors, or any other indicia of product identification identical or similar to, or identifiable with, those used for any brand of cigarettes or smokeless tobacco, *unless such athletic, musical, artistic, or other social or cultural event, or series of events, or any team or entry participating in such an event is an adult event or series of events.*
>
> Nothing, however, shall prevent a manufacturer, distributor, or retailer from sponsoring any athletic, musical, artistic, or other social or cultural event, *series of events*, entry or team in its corporate name.[269]

An "adult event" should be defined as "any event in which the number of persons under the age of eighteen in attendance and viewing on television is (1) less than fifteen percent of the total persons in attendance and viewing on television, or (2) less than two million." A "series of events" should be defined as "the sum of individual events which comprise a series or season." Some examples would include, but are not limited to, NASCAR auto racing, Major League Baseball, the National Basketball League, the National Football League, and the PGA tour. In determining whether a series of events is an "adult series of events," the total number of persons under the age of eighteen in attendance or viewing on television for the entire series or season should be used to arrive at a per event average. The per event average must satisfy the criteria for an "adult event" in order to be considered an "adult series of events."

The fifteen percent/two million person benchmark is borrowed from the FDA's method of determining what constitutes an adult publication. In arriving at fifteen percent, the FDA considered the percentage of children between the ages of five and seventeen, the ages of young readers, in the United States and found them to be roughly fifteen percent of the total population.[270] Thus, any magazine with a readership above this percentage would be more directed at children. This approach makes sense because any percentage higher than the total percentage of the child population would mean that a disproportionate number of children and adolescents were being exposed to tobacco advertising. Because the total percentage of persons under eighteen in the United States is

---

269. This is substantially the same language as the FDA's rule as codified at 21 CFR § 897.34(c) (1996). The additions included in this proposal have been italicized.

270. FDA Regulations, *supra* note 3, at 44,516.

25.7 percent, [271] the argument could be made that the percentage threshold should be twenty-five percent. However, that figure would allow brand-name advertising to reach a substantial number of children. Thus, fifteen percent is still a reasonable number for sponsored events.

The two million person benchmark set forth in the print regulations is also justified for sponsored events. Some deference should be afforded the FDA's conclusion that, at some point, the total number of children affected by tobacco advertisements becomes substantial.[272] Two million seems reasonable, and there is no reason to depart from that number for sponsored events.

The "adult series of events" definition is included as a matter of consistency and fairness. In NASCAR Winston Cup Racing, there are thirty-two events each year. The sponsorship of a racing team is usually a commitment that lasts the entire racing season. The determination of whether a tobacco company can sponsor a team in a brand-name should be made only once. It would be confusing to all involved to say that with respect to races A, B, and C, it is the NASCAR Winston Cup Racing Series, but at races X, Y, and Z, it is the NASCAR R.J. Reynolds Racing Series. Since all thirty-two events make up an entire NASCAR season, it is fair to both the FDA and the tobacco companies to make one computation and determine whether the series is eligible for brand-name sponsorships.

The figures used to determine brand-name sponsorship eligibility should be based on the last series or season, if available. If not, then the figures should be based on reliable estimations. Simmons Market Research Bureau, Inc., which the FDA cites with approval in the print advertising regulations and who currently provides marketing research for NASCAR, and Nielsen, which provides television ratings, are reliable sources who are readily available to determine if the event or series can be sponsored in a brand-name. The burden of verifying that the statistics are correct should be upon the tobacco industry. Furthermore, if brand-name sponsorship does not meet the criteria, the tobacco companies still have available the option to sponsor an event in its corporate name. Thus, this proposal does not close all avenues of advertisement if the fifteen percent/two million benchmark cannot be satisfied.

This proposal is a solution to the FDA's overly broad, existing rule. It promotes the FDA's substantial interest in protecting children and adolescents from tobacco advertising, yet it is consistent with commercial speech jurisprudence in that it is narrowly drawn. This proposal differentiates between events that are primarily adult and those that are not. This proposal should be upheld in court because it does not impose a blanket ban on advertisements, but rather sets forth criteria which must be satisfied for the sponsorship ban to take effect. Finally, the proposal is flexible. While the number of children that attend NASCAR races "may be growing," the fact remains that NASCAR is primarily an adult sport. As long as this is true, the sponsorship ban should not apply. When, and if, the number of children reach the benchmark numbers, the ban

---

271. U.S. Bureau of the Census, 1990 Census.

272. FDA Regulations, *supra* note 3, at 44,514.

would take effect and the FDA's interest would be served.

CONCLUSION

The consequences associated with tobacco use cannot be ignored. Hundreds of thousands of lives are lost each year because of tobacco use. Although tobacco is a legal product, the tobacco industry must take a responsible position in its efforts not to intentionally market its product to children and adolescents. Nevertheless, the concepts of individual liberty and freedom are principles upon which this nation was built. The First Amendment cannot be trampled upon solely because of what the government perceives as best for society. With respect to these competing interests, the need for an effective, but constitutional, solution is great. Adopting a fifteen percent/two million benchmark strikes a reasonable balance between furthering the FDA's interest in protecting future generations from the harms of tobacco use and preserving the integrity of the First Amendment.

UNITED STATES POSTAL SERVICE™

Statement of Ownership, Management, and Circulation

(Required by 39 USC 3685)

1. Publication Title	2. Publication Number	3. Filing Date
Indiana Law Review	0090_4198	9-18-97
4. Issue Frequency	**5. Number of Issues Published Annually**	**6. Annual Subscription Price**
Quarterly	Four	$25.00

7. Complete Mailing Address of Known Office of Publication *(Not printer) (Street, city, county, state, and ZIP+4)*

735 W. New York St., Indianapolis, Marion, Indiana 46202-5194

Contact Person: Chris Paynter

Telephone: (317) 274-4440

8. Complete Mailing Address of Headquarters or General Business Office of Publisher *(Not printer)*

735 W. New York St., Indianapolis, Indiana 46202-5194

9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor *(Do not leave blank)*

Publisher *(Name and complete mailing address)*

Indiana University School of Law - Indianapolis
735 W. New York St.
Indianapolis, Indiana 46202-5194

Editor *(Name and complete mailing address)*

Joseph Bowman
735 W. New York St.
Indianapolis, Indiana 46202-5194

Managing Editor *(Name and complete mailing address)*

Michael Dorelli
735 W. New York St.
Indianapolis, Indiana 46202-5194

10. Owner *(Do not leave blank. If the publication is owned by a corporation, give the name and address of the corporation immediately followed by the names and addresses of all stockholders owning or holding 1 percent or more of the total amount of stock. If not owned by a corporation, give the names and addresses of the individual owners. If owned by a partnership or other unincorporated firm, give its name and address as well as those of each individual owner. If the publication is published by a nonprofit organization, give its name and address.)*

Full Name	Complete Mailing Address
Indiana University Board of Trustees	Indiana University - Bloomington Indiana Memorial Union Room 005 Bloomington, Indiana 47405

11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. If none, check box → ☒ None

Full Name	Complete Mailing Address
NONE	NONE

12. Tax Status *(For completion by nonprofit organizations authorized to mail at special rates) (Check one)*

The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes:

☒ Has Not Changed During Preceding 12 Months

☐ Has Changed During Preceding 12 Months *(Publisher must submit explanation of change with this statement)*

13. Publication Title		14. Issue Date for Circulation Data Below	
Indiana Law Review		September 1997	
15. Extent and Nature of Circulation		Average No. Copies Each Issue During Preceding 12 Months	Actual No. Copies of Single Issue Published Nearest to Filing Date
a. Total Number of Copies *(Net press run)*		1345	1600
b. Paid and/or Requested Circulation	(1) Sales Through Dealers and Carriers, Street Vendors, and Counter Sales *(Not mailed)*	190	600
	(2) Paid or Requested Mail Subscriptions *(Include advertiser's proof copies and exchange copies)*	875	885
c. Total Paid and/or Requested Circulation *(Sum of 15b(1) and 15b(2))* ▶		1065	1485
d. Free Distribution by Mail *(Samples, complimentary, and other free)*		50	50
e. Free Distribution Outside the Mail *(Carriers or other means)*		45	30
f. Total Free Distribution *(Sum of 15d and 15e)* ▶		95	80
g. Total Distribution *(Sum of 15c and 15f)* ▶		1160	1565
h. Copies not Distributed	(1) Office Use, Leftovers, Spoiled	185	35
	(2) Returns from News Agents	0	0
i. Total *(Sum of 15g, 15h(1), and 15h(2))* ▶		1345	1600
Percent Paid and/or Requested Circulation *(15c / 15g x 100)*		92%	95%

16. Publication of Statement of Ownership
☒ Publication required. Will be printed in the 31:1 issue of this publication.
☐ Publication not required.

17. Signature and Title of Editor, Publisher, Business Manager, or Owner	Date
Joseph C. Bauman — Editor-in-Chief	9-19-97

I certify that all information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including multiple damages and civil penalties).

Instructions to Publishers

1. Complete and file one copy of this form with your postmaster annually on or before October 1. Keep a copy of the completed form for your records.
2. In cases where the stockholder or security holder is a trustee, include in items 10 and 11 the name of the person or corporation for whom the trustee is acting. Also include the names and addresses of individuals who are stockholders who own or hold 1 percent or more of the total amount of bonds, mortgages, or other securities of the publishing corporation. In item 11, if none, check the box. Use blank sheets if more space is required.
3. Be sure to furnish all circulation information called for in item 15. Free circulation must be shown in items 15d, e, and f.
4. If the publication had second-class authorization as a general or requester publication, this Statement of Ownership, Management, and Circulation must be published; it must be printed in any issue in October or, if the publication is not published during October, the first issue printed after October.
5. In item 16, indicate the date of the issue in which this Statement of Ownership will be published.
6. Item 17 must be signed.

Failure to file or publish a statement of ownership may lead to suspension of second-class authorization.


MAR 1998
RECEIVED
Office of the Chancellor
IUPUI